AD_____

Award Number: DAMD17-98-1-8045

TITLE: Improving Clinical Diagnosis Through Change Detection in
Mammography

PRINCIPAL INVESTIGATOR: Yue-Joseph Wang, Ph.D.

CONTRACTING ORGANIZATION: The Catholic University of America
Washington, DC 20064

REPORT DATE: September 2000

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are
those of the author(s) and should not be construed as an official
Department of the Army position, policy or decision unless so
designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE September 2000	3. REPORT TYPE AND DATES COVERED Annual Summary (1 Sep 99 - 31 Aug 00)

4. TITLE AND SUBTITLE
Improving Clinical Diagnosis Through Change Detection in Mammography

5. FUNDING NUMBERS
DAMD17-98-1-8045

6. AUTHOR(S)
Yue-Joseph Wang, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Catholic University of America
Washington, DC 20064

E-MAIL:
wang@pluto.ee.cua.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

Temporal change of mass lesions overtime is a key piece of information in computer-aided diagnosis of breast cancer and treatment monitoring. For a specific patient, change detection depends on the ability to align the images of the mammogram sequence to a common axis, and the ability to build up memory about the image scene overtime. The process of aligning images to a common axis is termed image registration. The image scene representation is called site model. In the second year of this project, we developed a novel registration technique to align temporal sequences of the same patient, to construct a scene memory or site model, with the ultimate goal of performing change detection.

We developed (1) a new hybrid registration algorithm aimed at the registration of non-rigid objects with minimal a prior knowledge; (2) a new change quantification metric based on the joint relative entropy between two images; (3) a patient specific site model concept to image-guided lesion monitoring; (4) a methodology to combine multiple transforms together to determine a composite image transform; and (5) an improved statistical segmentation algorithm for sequences of images.

14. SUBJECT TERMS
Breast Cancer, image registration, change detection, mass detection.

15. NUMBER OF PAGES
123

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

20010302 051

3. TABLE OF CONTENTS

4. INTRODUCTION

Temporal change of mass lesions overtime is a key piece of information in computer-aided diagnosis of breast cancer and treatment monitoring, the **purpose** of the project is to develop an automatic change detection method to quantitatively extract the clinically important changes of suspicious lesions, upgrade the existing CAD system, and thus improve the clinical diagnosis of breast cancer. We will build a site model for each individual patient for monitoring the breast tissue changes and extend our current research on image registration and change detection to the early detection of breast cancer. **Specific aims** include: 1) registration and segmentation of deformable breast tissue structures across a series of mammograms; 2) construction of a site model of the mammogram for individual patients showing the locations of regions of interest and associated diagnostic information; 3) identification of clinically significant changes in both global and local mass areas within the breast; and 4) integration and evaluation of the developed techniques with existing CAD prototype. At conclusion of this project, we anticipate **achieving** the following: 1) establish a reliable technique of monitoring breast tissue changes associated with cancerous masses; 2) deliver a CAD prototype that can incorporate tissue change information from additional mammograms; 3) evaluate the merit of combining change detection and CAD for improved clinical diagnosis using multiple mammograms; and 4) acquire the experience necessary to explore multimodality imaging for unified detection, diagnosis and treatment assessment of breast cancer.

5. BODY-Annual Summary

The **long-term** goal of this career development project is to develop image guided diagnosis methodology through change detection in mammogram sequences for breast cancer detection. The research requires the knowledge of image analysis, image registration, change quantification, and machine intelligence.

During the second year of this career development project, I have developed a close research collaboration with Dr. Matthew Freedman (Radiologist) and Dr. Ben Lo (Medical Physicist) at Georgetown University Medical Center. I have also developed a strategic collaboration with Dr. Robert Clarke (Department of Oncology) at the Lombardi Cancer Center. Through them, I have learned more about breast cancer at both **imaging** and **molecular** levels. I have been serving as a Panel Member for the Study Sections on both **Bioinformatics** and **Bioimaging** for the National Cancer Institute since 2000.

As the Director of Imaging and Intelligent Informatics (I3) Laboratory, I am currently serving as the major advisor to eight full-time graduate students specifically working on breast cancer research. I am also the Member, Technical Committee (TC) on Neural Networks for Signal Processing (NNSP), IEEE Signal Processing Society, since 1999; the Member, Program Committee, IEEE Workshop on Neural Networks for Signal Processing, Australia 2000; the Member, Technical Committee, IEEE Workshop on Multimedia Information Processing, Australia 2000; the Session Chair, International Joint Conference on Neural Networks, Washington, DC 1999.

I have been promoted to the rank of **Associate Professor** after four-year faculty service at CUA.

As the research accomplishments during the second year, I have first identified the following major research tasks:

1. Construct a patient specific site model based on the outcome of image analysis including objects, surface, boundaries, and control points, of the normal tissues and detected/suspected lesions. This will provide a mathematical model for (1) high accuracy change monitoring considering the patient variation and (2) effective data fusion incorporating prior/domain specific information.
2. Develop a multiple step algorithm for two-dimensional image registration of image sequence data sets. It consists of three major components: (1) principle axes registration (PAR), (2) site model support control feature alignment with localized PAR, and (3) deformable data matching via thin-plate spline (TSP) interpolation.
3. Apply new algorithm to perform change detection from a set of sequence images based on information theory, where the clinical objectives are lesion verification/detection, lesion localization, and change quantification.

Follow this plan, major research accomplishments include:

5.1 New hybrid image registration algorithm

We have implemented a new hybrid registration algorithm aimed at the registration of non-rigid objects with minimal a prior knowledge, in which we have developed a methodology to combine multiple transforms together to determine a statistically composite geometric transform. The purposed algorithm combines rigid and non-rigid techniques to accomplish the registration tasks.

Figure 1. Patient specific site model.

The algorithm consists of two steps an initial step (rigid transform) which performs multi-object PAR registration where object correspondence is assumed known, and a final step (non-rigid transform) that uses thin-plate spline (TPS) based mapping where control point correspondence is determined via a
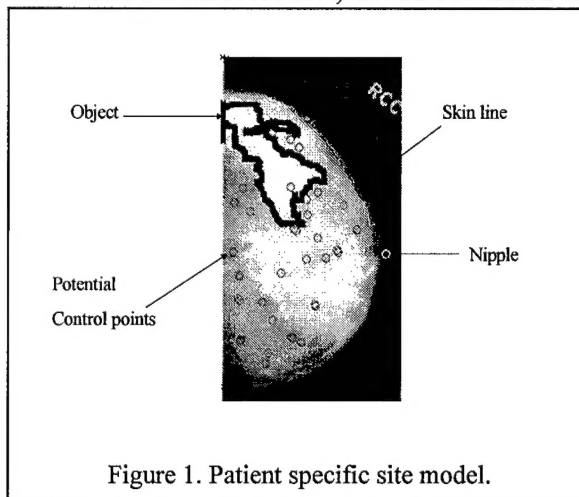
detection and correspondence algorithm. The combination of these two steps is new and provides many advantages over existing methods. The first advantage is no requirement for point correspondence in the initial step. Only object correspondence is required which is usually much easier computationally to determine. True point correspondence is required at some point in the processing, but performing the determination after the image has been preliminarily aligned should allow for a more focused or narrow control point search windows because potential control points should now be closer spatially. The second advantage is the ability to model non-rigid transforms by considering each rigid transform as a piece wise component of a total non-rigid transform similar to modeling a non-linear function by linear pieces. This approach is a departure from traditionally registration approaches which usually follow either rigid or non-rigid transforms. In particular, we apply the combination method to multiple PAR transforms, but the

Figure 2. Control point extraction.

method is generic and can be applied to any type of transform along as each cluster control point meets the particular requirement of the registration method in question. For example, to use an elastic registration method it is assumed we know the point correspondence of control points. In this algorithm, the image is assumed to contain several clustered control points, which follow a normal distribution, for which cluster correspondence is known (i.e. objects). The resulting transform now enables rigid transform methods to handle non-rigid transform assuming the clusters are sufficiently distributed through out image.
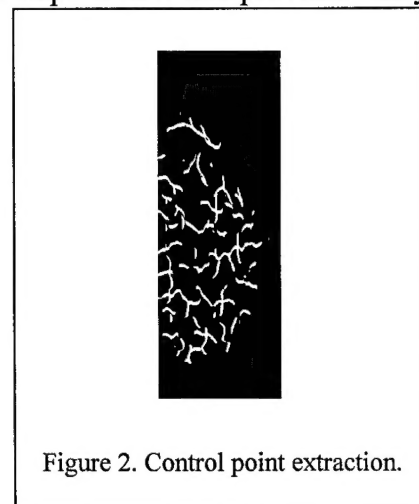
5.2 Construction of patient specific site model

We have developed a patient specific site model concept to image-guided lesion monitoring. The site model was developed to monitor a site from a sequence of aerial images. In medical imaging, the site model idea was modified to accomplish application such as lesion monitoring, and disease detection. In addition, through update procedures the site model allows for the

Figure 3. Control point matching.

examination of the entire sequence together, to show region progression or to further highlight small changes. The main modification to the site model idea was the creation of another variable to store changes. In traditional site model formulations, new objects are added back into the image, but in the medical environment the site image is untouched. The changes are stored in the change map. The site image is untouched because it forms the base frame for comparison so any modification could alter results.

5.3 New change quantification metric

We have developed a new change quantification metric based on the joint relative entropy between two images. Unlike other change detection metrics, the joint relative entropy is useful in detecting translation only changes. In addition, the results of the metric tell us how similar the blocks are to each other. Difference image analysis is also useful for translation change, but it is highly sensitive to

Figure 4. Control point matching over time.

noise and does not yield a measure of how close the blocks of data are to each other. In addition, this algorithm is used in the site model update to reprocess the segmented image given the images of the sequence. The major assumption is that the adjacent images contain the same view. This algorithm is
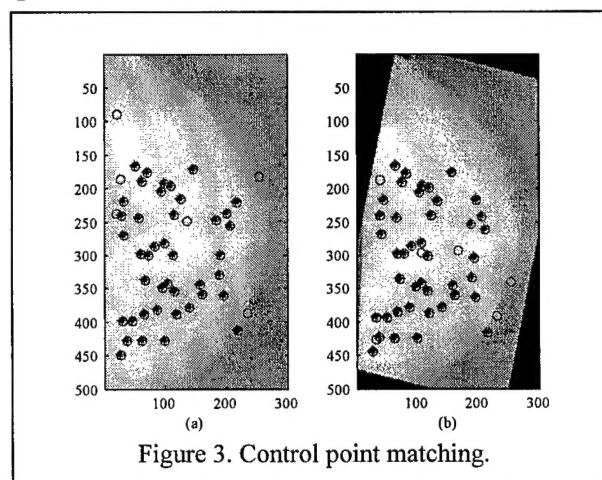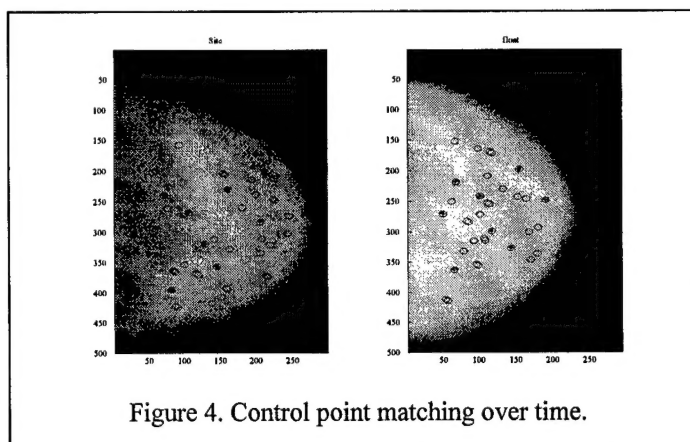
based on a 2D statistical segmentation algorithm where pixel relationship is assumed across adjacent pixels in the (x,y) direction. The algorithm extension takes advantage of the relationship between adjacent images. So, pixel neighborhood is considered in three directions (x,y,z). This additional information leads to a more robust segmentation for change detection.

5.4 Key results and discussion

The registration process is supported by the concept of a site model and site model operations. The site model is a mathematical representation of a scene under analysis. A basic site model contains a geometric description of a scenes objects (area, size, and other attributes), raw data, and simple user input (previous tumor locations). The environment interacts with the site model through the site model operations: construction, image-to-site registration and model parameter update. The site model is constructed by thoroughly processing the first image in the sequence to obtain the parameters. The site model supports registration in three main ways. First, the site model forms the reference frame (reference image) for all subsequent images, thus allowing all of the images in the sequence to be alignment to a common coordinate system. Second, the model stores registration parameters like object contours, control points, and user identified regions. This effectively integrates both manual and automatic control objects in a single place. Third, the model stores previously detected change, this enables the current registration process to exclude the previously detected changed portion from the current analysis which improves algorithm robustness. In this research, we focus on the rigid, affine, and polynomial based registration methods to register the sequence of mammograms of the same patient. Image-to-site model registration is performed by a multi-step algorithm consisting of an initial and final phase. The initial phase registers the images using the principle axis of the skin line in conjunction with segmented internal objects to form a multi-object global rigid spatial-coordinate transform followed by a simple look up table for the intensity transform. The final registration phase consists of a global thin-plate spline transform derived from the control points of the interior breast tissue.
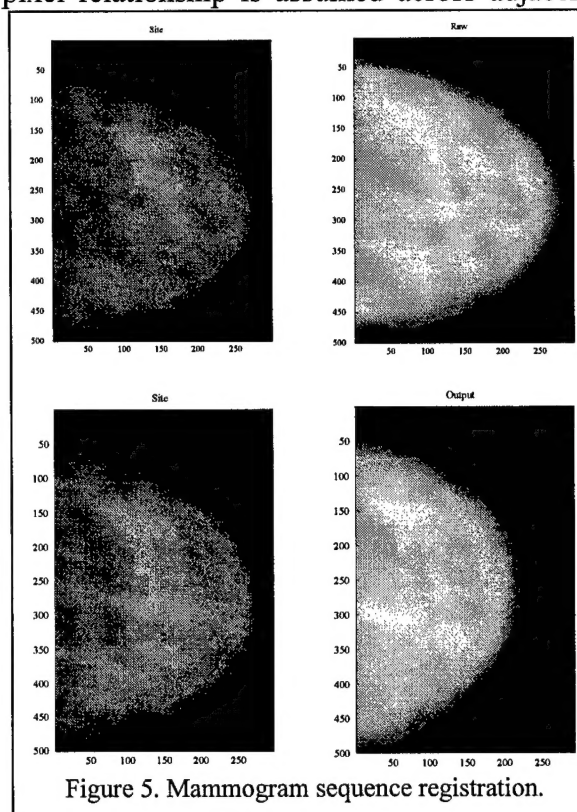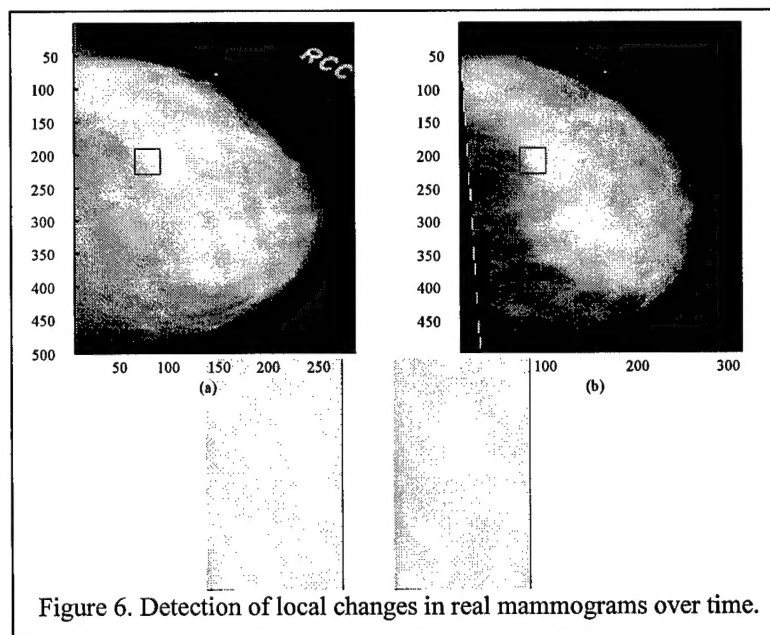

Figure 5. Mammogram sequence registration.

Figure 6. Detection of local changes in real mammograms over time.

Figure 1 shows the patient specific site model, where potential control points, skin line, nipple location, as well as object boundaries, for image registration, are extracted and stored. Figure 2 shows the result of control points extraction using our method. Figure 3 shows the corresponding control points in two similar breast phantoms. It can be seen that most control points are well matched using our PAR based initial registration. Figure 4 shows shows the corresponding control points in two real mammogram sequence. After our initial registration, stable control points are matched for further registration effort.

Figure 5 shows the results of two cases of mammogram registration using our hybrid algorithm. It can be seen that even with breast deformation, our method can find a good matching particularly for local changes. Figure 6 shows the result of combined image registration and local change detection. More results are provided in our attached Technical Report.
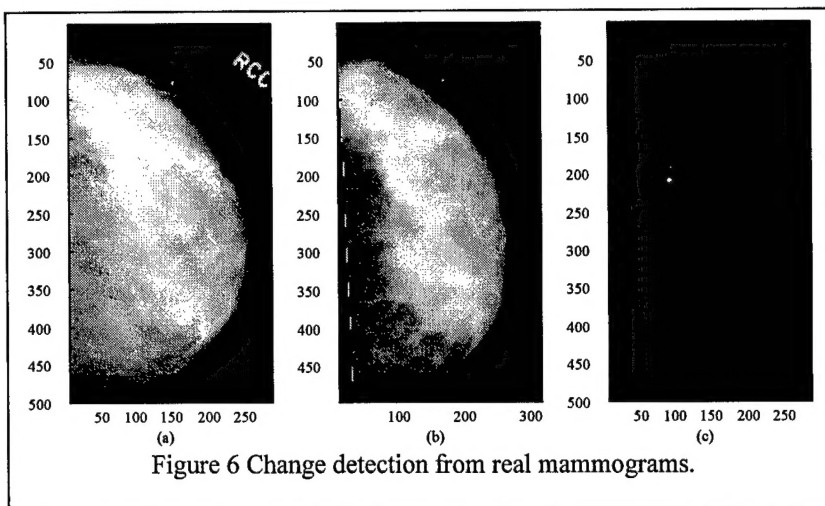

Figure 6 Change detection from real mammograms.

Change detection not only highlights existence of possible changed regions, but when combined with the site model provides a patient history by showing site progression. One of the key components of change detection is image registration. In this project, we applied our multi-step registration algorithm to mammogram sequences. Acceptable registration and change detection were obtained. Improvement in control object selection and control point extraction would go along way to improving the overall results. The key to registration is landmarks between the images. In this research, we use objects and points as landmarks. Current methods of object and point selection are image dependent and ad hoc. Incorrect assignment of control points/objects could cause erroneous transformation. This change detection is not exact, but would be sufficient to flag a radiologist to review the area. The main results of this study consisted of the automatic alignment of mammograms, detection of change in a local window, and implementation of a mechanism to store and build up patient information via the site model.

6. APPENDICES

6.1 Key Research Accomplishments

This report presents methodologies and techniques to aid in the automation of the change detection process. The change detection process finds application in medical imaging specifically applied to lesion diagnosis and tumor detection. This study is limited to determining change in a previously selected window (i.e. local change), not change on a global scale. This is accomplished by the development of site model supported change detection algorithm. The change detection algorithm is divided into four main tasks: site model construction, preprocessing, registration, and change detection and quantification. Site model construction and preprocessing use classical signal and image processing techniques to derive the site model parameters (i.e. build the model). Registration, the most challenging component, uses a novel multi-step algorithm consisting of multi-object principle axis registration (PAR) for initial registration and thin-plate spline (TPS) transformation of control points for final registration. Three methods for combining the multiple transforms of initial registration are considered. They are local, average, and finite mixture. Local combination yields images containing discontinuity on boundaries. Average combination produces a smooth image, but assumes a rigid transform for the rest of the image. Finite mixture combinations produces a smooth image and can be used to model non-rigid deformation with several rigid transforms. In this study, finite mixture is used because the breast is generally assumed to be a non-rigid body. The change detection phase is performed by a two step process. Step one compares the joint relative entropy of the two image blocks with a detection threshold. Step two combines object area and center of gravity differences between the blocks as a means of quantification.

This complete change detection algorithm was simulated with phantom images and real mammograms. The benefits of two steps in registration are apparent by looking at the mean square pixel error between no registration, single object PAR, and multi-PAR/TPS registration where the MSE drops almost 84% compared to only 70% with PAR alone. The change metric (joint global relative entropy (GRE)) was compared to two existing video sequence methods chi square and histogram difference. Joint GRE performed better as it was able to detect intensity changes, shift changes and shift/intensity changes. The quantification process estimated on average within 15% of the true objects size for the studies under considerations.

This complete process facilitates change detection by aligning the images and comparing corresponding regions of interest for change resulting in a accurate detection of local change and a patient specific site model showing image conditions over time. A key factor that governors this process is the alignment of the incoming mammograms to the site model. This process could be improved with more robust control object selection and control point selection, and obtaining sufficient distribution of control objects/points during the registration phase. Also, improvement of change quantification methods to consider more complex methods of description and analysis should result in more robust quantification.

6.2 Reportable Outcomes

1) H. Li, Y. Wang, K-J R. Liu, S-H B. Lo, and M. T. Freedman, "Computerized Radiographic Mass Detection I: Lesion Site Selection by Morphological Enhancement and Contextual Segmentation," to appear *IEEE Transactions on Medical Imaging* 2000.

2) H. Li, Y. Wang, K-J R. Liu, S-H B. Lo, and M. T. Freedman, "Computerized Radiographic Mass Detection II: Decision Support by Feature Database Visualization and Modular Neural Networks," to appear *IEEE Transactions on Medical Imaging* 2000.

3) J. Xuan, T. Adali, Y. Wang, and E. Siegel, "Automatic Detection of Foreign Objects in Computed Radiography," to appear *SPIE Journal of Biomedical Optics*, October 2000.

4) Y. Wang, L. Luo, M. T. Freedman, and S. Y. Kung, "Probabilistic Principal Component Subspaces: A Hierarchical Finite Mixture Model for Data Visualization," IEEE Transactions on Neural Networks, Vol. 11, No. 3, pp. 625-636, May 2000.

5) Kelvin Woods (Major Advisor: Dr. Yue Wang), Doctoral Dissertation, *Image Guided Diagnosis through Change Detection in Image Sequences*, The Catholic University of America, 2000.

6) K. Woods, Y. Wang, and M. T. Freedman, "Patient site model supported change detection," *Proc. SPIE Medical Imaging*, San Diego, Feb. 2000.

7) K. Woods and Y. Wang, "Patient site model supported change detection in mammogram sequences," (Abstract), Poster Presentation, Era of Hope 2000, Department of Defense.

The Catholic University of America

Technical Report

Improving Clinical Diagnosis of Breast Cancer through Change Detection in Mammogram Sequences

Kelvin Woods and Yue Wang

August 2000

Abstract

This report presents methodologies and techniques to aid in the automation of the change detection process. The change detection process finds application in medical imaging specifically applied to lesion diagnosis and tumor detection. This study is limited to determining change in a previously selected window (i.e. local change), not change on a global scale. This is accomplished by the development of site model supported change detection algorithm. The change detection algorithm is divided into four main tasks: site model construction, preprocessing, registration, and change detection/quantification. Site model construction and preprocessing use classical signal and image processing techniques to derive the site model parameters (i.e. build the model). Registration, the most challenging component, uses a novel multi-step algorithm consisting of multi-object principle axis registration (PAR) for initial registration and thin-plate spline (TPS) transformation of control points for final registration. Three methods for combining the multiple transforms of initial registration are considered. They are local, average, and finite mixture. Local combination yields images containing discontinuity on boundaries. Average combination produces a smooth image, but assumes a rigid transform for the rest of the image. Finite mixture combinations produces a smooth image and can be used to model non-rigid deformation with several rigid transforms. In this study, finite mixture is used because the breast is generally assumed to be a non-rigid body. The change detection phase is performed by a two step process. Step one compares the joint relative entropy of the two image blocks with a detection threshold. Step two combines object area and center of gravity differences between the blocks as a means of quantification.

This complete change detection algorithm was simulated with phantom images and real mammograms. The benefits of two steps in registration are apparent by looking at the mean square pixel error between no registration, single object PAR, and multi-PAR/TPS registration where the mse drops almost 84% compared to only 70% with PAR alone. The change metric (joint global relative entropy (GRE)) was compared to two existing video sequence methods chi square and histogram difference. Joint GRE performed better as it was able to detect intensity changes, shift changes and shift/intensity changes. The quantification process estimated on average within 15% of the true objects size for the studies under considerations.

This complete process facilitates change detection by aligning the images and comparing corresponding regions of interest for change resulting in a accurate detection of local change and a patient specific site model showing image conditions over time. A key factor that governors this process is the alignment of the incoming mammograms to the site model. This process could be improved with more robust control object selection and control point selection, and obtaining sufficient distribution of control objects/points during the registration phase. Also, improvement of change quantification methods to consider more complex methods of description and analysis should result in more robust quantification.

1

Chapter 1

Introduction

1.1 Background

Breast cancer is one of the leading causes of death among women today. To help combat this problem doctors use medical imaging (mammography) as a mechanism to screen patients and identify cases where further analysis is required. In breast cancer diagnosis, the mammography has proven to be the only way to detect cancer at its earliest stages, thus improving the patient survival probability[4]. A patient's survival probability is directly linked to tumor size upon detection. Tumor size has an apparent relationship to tumor grade or disease progression which can dictate treatment options. Studies have shown that women at age 40 and up are most at risk for developing breast cancer. Although this factor alone is not the sole contributor, most women over 40 have screening mammograms performed periodically (usually one or two years apart) in an effort to detect the existence or onset of a cancerous condition in the breast. This type of study is called breast cancer screening and usually is limited to asymptomatic women where craniocaudal (CC) and mediatorial oblique (MLO) mammographic views are acquired and analyzed for signs of cancer[4]. These images are reviewed manually by a radiologist following a prescribed procedure which specific viewing apparatus, lighting requirements, and amount of time per case [4]. Generally, a radiologist reviews four images of a single view (either CC or MLO) simultaneously. The images are the current left and right breast aligned over top of the left and right breast taken previously. Figure 1.1 shows the layout for the screening case. By aligning the images in this manner, change (tissue change) over time can better be identified. This tissue is a key indicator to the onset of a cancerous condition. Studies have shown a correspondence between tissue change and underlying biological change. This change is important for applications such as treatment monitoring and lesion diagnosis. Once change has been detected, further analysis of the region is performed.

1.2 Statement of Problem

Due to limited resources, radiologist often must review a massive number of cases during a period. Also, the constrains on resources have caused radiologist with less experience in mammography analysis to review cases. The review of this massive volume (around 8 images per case) of data and inexperience could cause missed tumors, delayed detection, and false positives which ultimately cause a reduced life expectation upon detection, unnecessary patient call backs, and unneeded needle biopsies.

To reduce some of the load on the radiologist and to improve diagnosis accuracy, development of automatic computer aided diagnosis (CAD) system for change detection have been explored [5], [6], [69]. These systems aim to automate portions of the analysis process. In order to accomplish this task, one must roughly model the analysis task performed by the radiologists in the course of an examination. Since this research focuses on change detection, the task modeling discussed here focuses on that task. The radiologists's analysis process consists of the following steps: (1) Acquire mammograms of previous and current visit; (2) Mount the image in specific order (see Figure 1.1); (3) Mentally examine images for similar landmarks and mentally adjust view; (4) Identifying corresponding regions and compare for change. From the examination of these four tasks, it is apparent that steps three and four would stand to benefit the most from automation as steps one and two are relatively simple.

Several key issues make automation of steps three and four extremely difficult, with step three being the most difficult. The issue is the fact that mammograms are complex images that do not contain any clearly defined landmarks. Secondly, differences in breast positioning and compression during acquisition could cause images of one scene to visually appear different. Finally, breast sizes and consistency can vary with time (e.g. weight loss, surgery, and age). The research of the clinical problem of change detection in a mammogram sequence of a single patient uncovers serval difficulties and complex technical problems. The first problem is how do you align a generally non-rigid object without apparent control points or landmarks? This problem is classified as a image registration
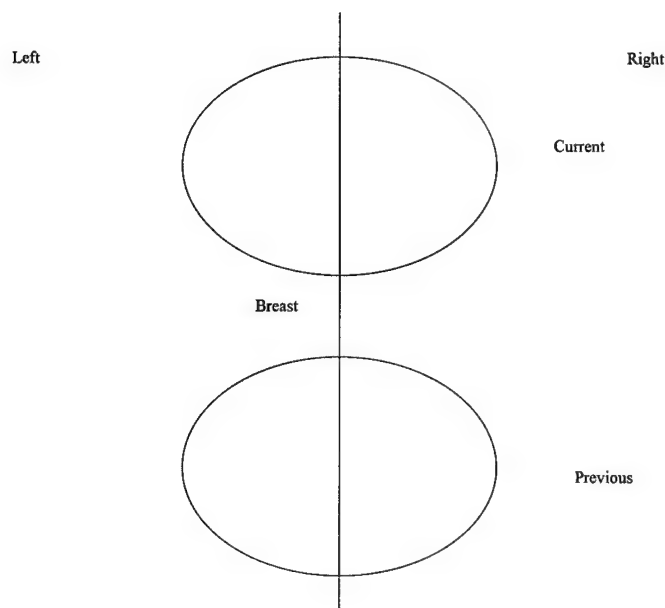
Figure 1.1: Layout of Screen mammogram analysis

problem. Image registration has been the topic of much research over the years [73]. The other problems are directly related to change detection. In mammograms, drastically different images can be attained from the same patient imaged at temporal displayed times. The key questions here are how do we discriminate natural change from cancerous change and how do we determine the type of change that has occurred? Often in medical imaging, the type of change that has occurred can direct the type of treatment required. Prime examples are treatment monitoring and tumor detections. This process we define as change quantification. This definition was motivated by the work of [18] where quantification is used to define the process of describing the image with some model parameters. The specific aim of this research is to study image registration and change detection to address the clinical and technical problems discussed above. The result will be a semi-automatic change detection algorithm.

1.3 Technical Review

Two main approaches were developed to deal with the problem of automatic change detection in mammograms. They are approaches based on processing a single view of a single breast [5], [6] and approaches based on single view of multiple breasts (left and right) [8]. [69] presented work that developed an approach to consider both single and multiple view processing. Use of multiple breast views leads to additional problems because women typically have significantly different structures between left and right breasts [1]. This causes natural asymmetry to be flagged as change or lead to landmark confusion [5] while single breast approaches do not have the problem of dealing with asymmetry. So, most of the research attention has been focused on single breast approaches. Generally, single breast approaches contain three main steps: (1) preprocessing of the images searching for control points or regions for use in registration, (2) registration, to align the images into a common framework, and (3) detection and analysis of local change. The preprocessing is generally handled by classical image processing techniques such as segmentation, morphological filtering, edge detection, and feature extraction. The registration process is performed by both rigid and non-rigid forms, but generally the breast is considered a deformable object thus non-rigid forms of registration should be used [73]. Finally, the local change analysis is performed with various techniques ranging in complexity from difference image analysis [15] to principle component analysis [81].

Three main research groups have attempted to address the problems of mammogram registration and change detection. Group [5] approached these problems with a two layered approach. In their approach, they perform a sequence of two polynomial based (thin-plate spline TPS) registration using different sets of control points. The first set of control points were extracted from the smoothed dense tissue boundary (i.e. brightest region on mammogram). The second set was extracted from the interior region of the dense tissue. Correspondence between control points for the first transform was performed by matching points on the reference image contour with similar points on the float image contour with the same maximum curvature. For the second transform, points with matching LAWS's texture features [87] were matched as control points. This approach has problems when the dense tissue does not

occupy a large percentage of the breast which typically occurs in radio-lucent breast [1, Breast book]. In cases like this, error occurs in transforms when the point to be transformed is far away from the control points thus reducing the effectiveness of the control points.

Another approach to mammogram registration and change detection was developed by [6]. They consider these problems by asserting that accurate registration of mammograms is intractable except with elastic transforms, and the only solution is regional registration [7]. In regional registration, localized areas of the two mammograms are aligned based on their distance from control points. In their approach, monotony operators are used to extract vertical and horizontal elongated structures (milk ducts, and blood vessels) in the image which they assume to be generally stable between images in the sequence. A three-pass Gaussian filter is used on the original mammogram to mask less prominent structures. This reduces the complexity and limits the monotony operators to detecting the dominate structures. The cross points of these horizontal and vertical structures make up the pool of potential control points. Correspondence between the current image control points and reference image control points is accomplished by comparing the respective control point signatures. The signatures are created by counting the number of non-zero pixels that lie in a rectangle that is rotated around the control point. In this configuration, the direction of the longest structure would yield the highest valve in the signature. The similarity of the signatures is used as the matching criteria. These values are then passed into a thresholded accumulator matrix for final point selection. To localize the area where signatures are compared, the nipple location in both images is used to determine a neighborhood region that surrounds the potential control point. This reduces processing and decreases the probability of false alarm. Using these control points, regions (of any shape) are determined on the current image by calculating the distance from a subset of the detected control points. Finally, the regions are compared for change. This method overcomes the erroneous interpolation problem experienced by [5], but the algorithm uses ad hoc point matching criteria, localize window size selection, and threshold determination. In addition, [7] assumes a small mis-registration that restricts the generality of this approach. Both [5] and [6] mainly address registration so, simple change detection methodologies based on difference image analysis and wavelets respectively are used for their change analysis. In [9]'s approach, the registration is performed by a radial basis function (RBF) interpolation process. This approach as other in polynomial based registration methods depends heavily on the existence of control points in the image pair. This approach only uses control points on the skin line of the breast which has been extracted through threshold based image segmentation. Control point correspondence is obtained by finding contour points that are equidistant (measured in the number of contour points from the corresponding nipple) from the nipple. The control points are then used to solve for RBF parameters which yield the desired transform. Since the control point are selected only from the skin line, internal structures are not considered in registration. Thus, this method is unable to track non-rigid changes that occur inside the breast. In addition, use of threshold based segmentation could lead to a noisy contour.

Although these methods have had success on limited databases, their limitations could cause erroneous results when examining mammograms in a more general sense. For instance, consider a mammogram sequence where both images contain a small dense tissue area (relative to total breast tissue size). Using [63], the control points would be clustered around the dense tissue area leaving the rest of the image not modeled. So, any transform derived from these points could not accurately capture any deformation in the not modeled portion of the image thus causing mis-registration. In addition, consider that the same sequence has a large initial misalignment. This causes the window sizes, thresholds, and signature matching criteria of [6] to be manually modified to correctly process. The approach [69] is insensitive to the above conditions, but would not accurately model the internal structures because no control points exist in that region. This short fall could possibly cause the detection of false or missed change. The limitations of [5] [6] [69] are listed in Table 1.1.

Another problem not considered by the above three approaches is a sequences containing more than two images (i.e. I_i, I_{i-1}, I_{i-2},......). Sometimes in medical analysis, the radiologist will examine further back than previous images as some change can only be seen over a longer periods of time. In satellite imaging, site monitoring is a similar task. In this task, sites are monitored through several images (generally two or more). To accomplish this task [79] uses the site model. The site model is a multimedia representation of an image scene to include object shapes location, segmented version of scene, previous location of change, extracted features, and a prior domain expert information. Through the site model operations of construction, registration, and update the site model tracks the scene over time. This same approach could be used to analyze an anatomical region such as the breast, brain,or prostate in temporal studies.

1.4   Approach

Thus, considering the limitations listed in Table 1.1 and site model theories, a new algorithm is proposed to perform non-rigid registration applied to a mammogram sequence. In this algorithm the registration is perform in two steps. The first step is called initial registration and it aims to correct large global misalignment by treating the breast as a sum of rigid objects and performing a multi-object principle axis registration(PAR). The objects include large

Limitations	Effect of Limitations
Wirth Method	
Only use control points on the skin-line.	Unable to consider deformation of internal structures
Number of contour points between control points as measure of control point matching.	Assumes that the number contour points between two control points is constant across the float and reference image.
Difference image analysis (detection only).	No quantification
Sallam Method	
Used the boundary and interior of dense tissue to determine control points.	Control points do not model complete image deformation in case when dense tissue is a small percentage of image
Used threshold methods to segment image.	Yields different contours if intensity ranges differ for reference and float image.
Difference image analysis (detection only)	No quantification
Brzakovic Method	
Assume small initial mis-registration.	Limits use to cases of small registration.
Image dependent processing parameters such as signature search wintow, size of monotony operators, and thresholds.	Requires new parameters for each image.
Histogram analysis using raw images (detection only).	No quantification
Adhoc signature matching method.	Assumes the longest arm of signature will remain the same in float and reference images.

Table 1.1: Limitation of existing Mammogram registration algorithms

clustering of similar tissue types and the breast skin line. An individual PAR transform is calculated for each object. Each pixel x_i is then passed through each of the T_k transforms resulting in multiple point matching $\acute{x}_{ik}$ in the new image. The final point location $\acute{x}_i$ is formed by weighting each point $\acute{x}_{ik}$ by the probability z_{ik} that the point x_i was transformed by T_k (or probability that x_i belongs to class k). z_{ik} is derived by considering each of the objects as a cluster of control points described by a normal distribution. Thus similar [19], we assume that each (x, y) locations to be made up of a sum of these normal distribution which can be modeled as a finite mixture.

This formulation allows for a weighting of the transform T_k to determine the final transform T. Thus, creating a global interpolative transform that weights local characteristics based on their probability of membership. The next step in the registration process is called final registration. In this step, non-rigid displacements between images are accounted for using a polynomial based (thin-plate spline) registration. Polynomial based algorithms depend heavily on the existence of control points between the images. To obtain the control points, we follow a modified version of the approach discussed in [7] which it extracts the elongated structure from the mammogram and uses the cross points of vertical and horizontal structures as the control points. The approach is modified by using the Pearson correlation coefficient [14] to match the potential control point signatures instead of the direction of the longest arm of the signature.

Similar to registration, change detection is performed by a two step process. The process consists of a detection phase and quantification phase. The detection phase consists of measuring the relative entropy between the joint histogram of the float and reference images with the joint histogram of reference with itself. The quantification phase uses basic geometry to determine an object's area and center of gravity which are then compared to determine if the object has change. To add the ability to study longer sequences, the site model was used to support the registration and change detection process. The site model supports the registration process by defining a reference frame which all subsequent images will be registered. The site model also fuses user input knowledge with automatically extracted data into a single model to be used in the registration process. As for change detection the site model stores the detected changes along with site memory and any other parameter updates.

The automatic change detection algorithm can be summarized into three main steps as outlined below.

Initial Registration

    Preprocess mammogram for skin line and internal objects.

    Use multi-object PAR on breast tissue using the skin line and internal object to form a finite cluster transform.

Final Registration.

    Preprocess the PAR transformed image searching for control points and transform coefficients.
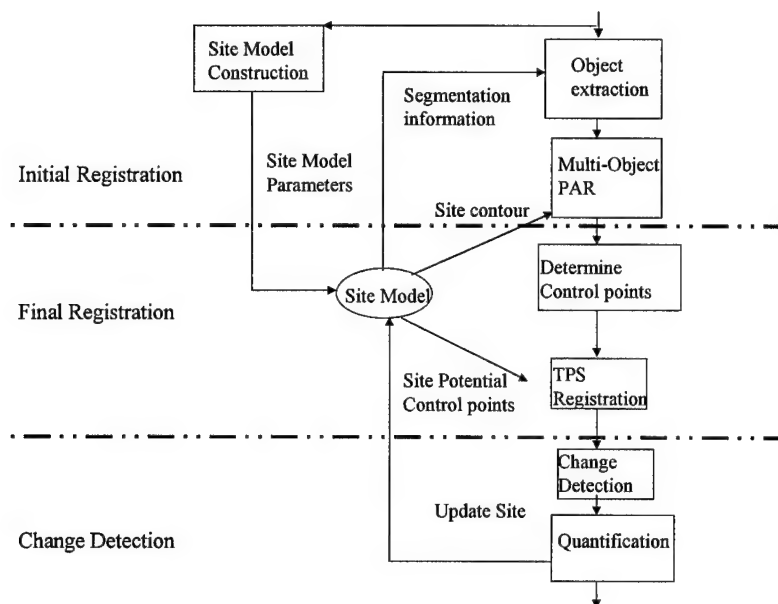
Figure 1.2: Change detection processing system flow

Use TPS formulation to determine the required transform.

Change Analysis

Use relative entropy for change detection criteria between the image blocks.

Quantify change by determining difference in object area and center of gravity.

Update change map located in the site model.

A complete flow diagram of the process is shown in Figure 1.2.

1.5   Research Scope

During the development of this algorithm, several assumptions were made in order to bound the scope of this research. First, the mammograms are assumed to be CC and MLO views only (i.e. screening mammograms) of the same patient acquired overtime. Second, the radiologist initializes the site model parameters by identifying areas of interest (local change windows) and other prominent landmark points (calcifications, large blood vessels) in the first image of the sequence. Third, the type of change was limited to growth of a mass, or shrinkage of a mass. Microcalcification changes can be detected, but will not be considered because drastic gray level difference between microcalcifications and non microcalcifications. Although, if present in both images of the sequence they may be used as control points. Fourth, the amount of initial mis-registration is bounded so the skin lines of each breast are not more that ±25°rotated from each other.

1.6   Contributions

The pursuit of this research has led to several contributions in image processing and medical imaging. Contribution one is the development of a new hybrid registration algorithm aimed at the registration of non-rigid objects with minimal a pori knowledge. Usually, non-rigid objects are registered with elastic or deformable methods which require knowledge of a sufficient number of control point pairs. While some rigid methods relax this requirement and usually only require object correspondence, for example, surface matching and principle axis methods. Use of rigid methods alone, in non-rigid problems, would allow for limited correspondence knowledge, but could not accurately model expected non-rigid deformations. The purposed algorithm combines rigid and non-rigid techniques to accomplish the registration tasks. The algorithm consists of two steps an initial step (rigid transform) which preforms multi-object PAR registration where object correspondence is assumed known, and a final step (non-rigid transform) that uses thin-plate spline (TPS) based mapping where control point correspondence is determined via a detection and correspondence algorithm. The combination of these two steps is new and provides many advantages over existing methods. The first advantage is no requirement for point correspondence in the initial step. Only object

correspondence is required which is usually much easier computationally to determine. True, point correspondence is required at some point in the processing, but performing the determination after the image has been preliminarily aligned should allow for a more focused or narrow control point search windows because potential control points should now be closer spatially. The second advantage is the ability to model non-rigid transforms by considering each rigid transform as a piece wise component of a total non-rigid transform similar to modeling a non-linear function by linear pieces[77]. This approach is a departure from traditionally registration approaches which usually follow either rigid or non-rigid transforms[73].

Contribution two is the development of a new change metric based on the joint relative entropy between two images. Unlike other change detection metrics [10], the joint relative entropy is useful in detecting translation only changes. In addition, the result of the metric tell us how similar the blocks are to each other. Difference image analysis is also useful for translation change, but it is highly sensitive to noise and does not yield a measure of how close the blocks of data are to each other.

Contribution three is the application of the site model concept to medical imaging. The site model was develop to monitor a site from a sequence of aerial images [13]. In medical imaging, the site model idea was modified to accomplish application such as lesion monitoring, and disease detection. In addition, through update procedures the site model allows for the examination of the entire sequence together, to show region progression or to further highlight small changes. The main modification to the site model idea was the creation of another variable to store changes. In traditional site model formulations, new objects are added back into the image, but in the medical environment the site image is untouched. The changes are stored in the change map. The site image is untouched because it forms the base frame for comparison so any modification could alter results.

Contribution four is the development of a methodology to combine multiple transforms together to determine a composite image transform. In this research, we apply the combination method to multiple PAR transforms, but the method is generic and can be applied to any type of transform along as each cluster control point meets the particular requirement of the registration method in question. For example, to use an elastic registration method it is assumed we know the point correspondence of control points. In this algorithm, the image is assumed to contain several clustered control points, which follow a normal distribution, for which cluster correspondence is known (i.e. objects). The resulting transform now enables rigid transform methods to handle non-rigid transform assuming the clusters are sufficiently distributed through out image.

Contribution Five is the development of a new statistical segmentation algorithm for sequences of images. This algorithm is used in the site model update to reprocess the segmented image given the images of the sequence. The major assumption is that the adjacent images contain the same view. This algorithm is based on a 2D statistical segmentation algorithm where pixel relationship is assumed across adjacent pixels in the (x, y) direction. The algorithm extension takes advantage of the relationship between adjacent images. So, pixel neighborhood is considered in three directions (x, y, z). This additional information leads to a more robust segmentation as seen in [54].

1.7   report Organization

This report is organized into seven chapters. The first chapter contains an introduction, background, problem statement, and contributions. The second chapter gives a brief tutorial on mammogram formation and screening procedures. Chapter three discusses the algorithms involved in the site model construction and update. Followed by chapter four that contains the techniques for image-to-site model registration. Chapter five discusses change detection while chapter six presents and discuss global results. Finally, chapter seven presents future research direction.

Chapter 2

Mammography formation and Screening

2.1 Introduction

Breast cancer is one of the leading causes of cancer related deaths among women. Each year more than 100,000 cases are diagnosed and more than 40,000 women die[1]. For many years researchers have studied breast cancer in search of an understand of breast cancer development. A high prediction rate of who will develop breast cancer is still an impossible task, although several factors have been identified as leading to the increase risk of breast cancer development. These factors include: gender, age, family history, age of first-term pregnancy, and previous history of breast cancer. Because of the gender factor, all women are at risk of developing breast cancer. In fact, women are 100 times more likely of developing breast cancer than men [4]. Breast cancer is a progressive disease, evolving through stages of growth. The size of the tumor size when detect has an apparent relationship to tumor grade and should be considered an important prognostic factor. Mammography, a form of X-ray imaging, has been shown to be the only method currently available for the reliable detection of early, non-palable, and potentially curable breast cancer [3]. So, women starting around the age of 40 are imaged every two years or so. These mammograms are put through rigorous examination for possible cancerous regions utilizing a process called screening mammogram. The rest of this chapter is organized as follows: tutorial on mammogram formations, and explanation of screening mammogram process.

2.2 Mammogram Formation

Mammography is an X-ray image of the breast used to detect, diagnose, or monitor cancerous conditions. It is usually performed by a trained technician with the ultimate goal of imaging as much breast tissue as possible. The patient is usually standing with her breast compressed against a support plate [2]. Compression of the breast is performed to equalize the thickness across the breast which produces a uniform image. A mammogram system is generally composed of four main components: X-Ray generator, compression device, scatter grid, and acquisition hardware. The general mammogram process is defined by these four steps. (1) arrange the breast in the compression apparatus. (2) Transmit a given X-Ray spectrum through the tissue. (3) Collect the X-rays and calculate the signal strength. (4) Form image using the results form in step (3). Figure 2.1 shows the arrangement of the components in relation to the breast to be imaged. The usability of the images is directly dependent on the image quality. Image quality is effected by several interrelated factors such as: contrast, which is useful in soft tissue examination; unsharpness, which is useful for small calcification; amount of X-Rays absorbed by breast tissue, where higher level increase contrast but put the patient at risk for radiation-induced carcinogenesis [4]; and high dynamic range which handles variation of the transmission over the entire mammogram. Thus, the goal is to determine compromises that best match the given factors. Next, each of the components in the Figure 2.1 will be discussed in more detail.

X-rays are produced by energy conversion when high speed electrons from the cathode hit the anode target as shown in Figure 2.2. The electrons are discharged from the cathode as a result of heating. This discharge is called therminic emission. X-rays (photons) are created when the electrons hit the atoms present in the anode. The area of the anode that is bombarded by the electrons is called the focal spot. The focal spot is directly related to image resolution. The smaller the focal spot the better the resolution. Since the X-ray emission from the anode is isotropic, shielding is needed to reduce undesired exposures to the patient and film. The shielding is performed by an elongated tube with a single opening. The tube opening is capped with a collimator to further reduce unwanted radiation emission.

The radiation is composed of three general energy levels low, medium, and high. The low and high energy photons are filtered out because low level photons usually are attenuated some much by the tissue that they do not
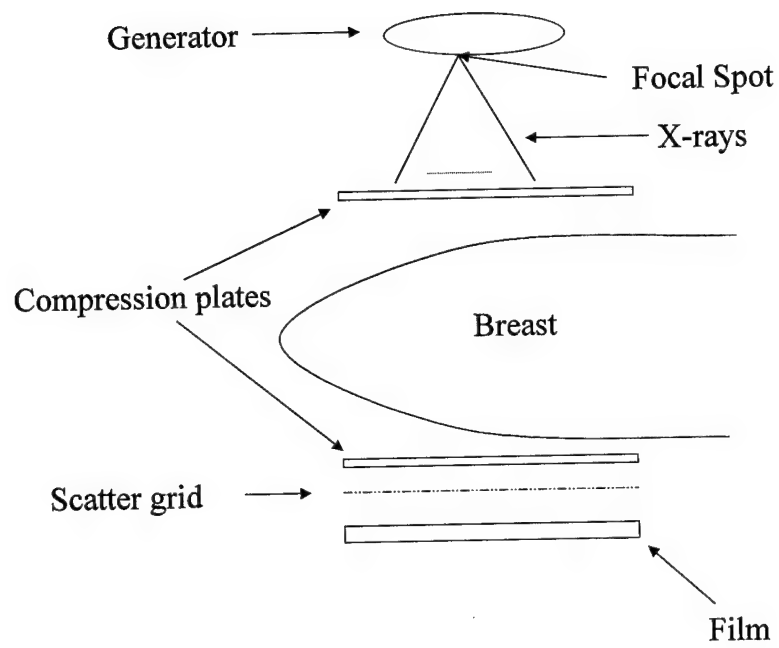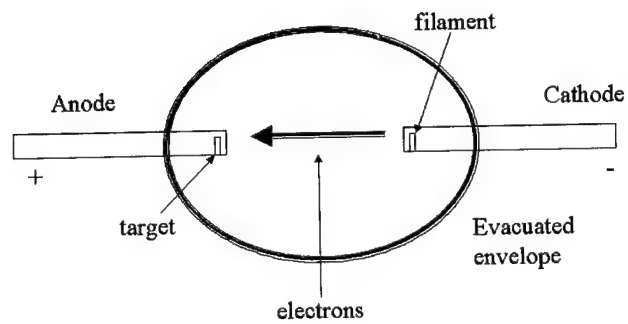
Figure 2.1: Mammogram System

Figure 2.2: Major components of X-ray

reach the film and the high level photons are unchanged by the tissue causing a low contrast image. This filter is used to shape the spectrum to achieve the best image quality. Most frequently, a molybolemun filter is used, but this is variable based on breast composition and thickness. Breast tissue composition goes through several phases of development during a womens' life. In each of these stages the breast can be composed of different tissue types. For example, in infancy the breast is mostly composed of adipose tissue while in puberty the fibroglandular tissue develops, and in maturity the fibroglandular tissue is replaced by fat tissue. Each of these tissue types attenuate the X-rays differently which yields different absorption rates.

The next component of a mammogram system is breast compression. Compression is performed using flat compression plates. A main advantage to compression is the breast tissue is forced to a uniform thickness. This avoids the problem of overexposing the thinner regions (near nipple) and underexposing the thicker regions (near chest wall). A second advantage is that the compression holds the breast in place during imaging. This reduces image unsharpness caused from tissue motion. Other advantages of compression are reduced absorption rates because the breast tissue is now thinner, shorter exposure time because the x-ray have a shorter distance to travel, and confusing and overlapping structures are separated.

Following the breast compression is the scatter grid. The scatter grid is designed to drastically attenuate the photons that are hitting the plate obliquely. These photons are more than likely the result of scattering from within the breast tissue. Scatter grids are composed of thin strips of metal laid with a particular spacing. Grids come in variety of different configurations. They are measured using a term a called grid ratio. This is defined as the ratio of the length to strip spacing. When the scattered photons are removed there is an increase the image contrast. In [2] contrast was improved by 17%, 37%, and 54% with the use of filters with ratio valves of 2,4, and 8.

The final component of Figure 2.1 is acquisition hardware. Acquisition hardware includes the process that receives the photons from the scatter grid and then translates it onto the film. This process contains two major steps. The first step converts the photon into visible spectrum by exposing a luminescent intensifying screen to the photons. This reaction produces light which is then used to expose film and form the radiographic image. Next, this image is transformed into a visible image by standard developing techniques.

2.3   Mammogram Screening

Screening mammograms is the term given to the periodic mammograms used in early detection of possible cancers conditions. The question the radiologist wants to answer using mammograms is, "Is this mammogram completely normal or is addition analysis required?" The major goal of mammography is to image the breast in order to detect cancerous conditions at its earliest stages. With this goal in mind technicians generally try to arrange the breast to image as much of the tissue as possible. Since the breast is a three dimensional organ, it is important to obtain multiple views so confusing or overlapping structures can be resolved. Generally, in screening studies the mediolateral oblique (MLO) and craniocaudal (CC) projects are obtained [1]. Together these two projections visualize the majority of the breast tissue, although, if sufficient compression is not achieved then the deep tissue close to the chest wall will not be imaged. Figure 2.3 and Figure 2.4 shows examples of CC and MLO compression views with a corresponding mammogram.

The mediolateral oblique projection is the most useful projection because this view projects most of the breast tissue onto the image including breast tissue close to the chest wall. In this projection, the compression plane is oblique not the patient. The compression plane extends through the nipple from the upper outer quadrant of the breast to the lower inner quadrant of the breast as shown in Figure 2.4. On the other hand, in the craniocaudal projection the compression plane is perpendicular to the chest wall. This view shows the thinner portion of the breast, but can often miss the thicker portion because of positioning. Usually, after the MLO and CC views have been examined, additional views may be required depending on the review results. The other supplement views include: lateral, medial, lateromedial, and straight mediolateral. Use of these views depends heavily on the particular cancerous sign.
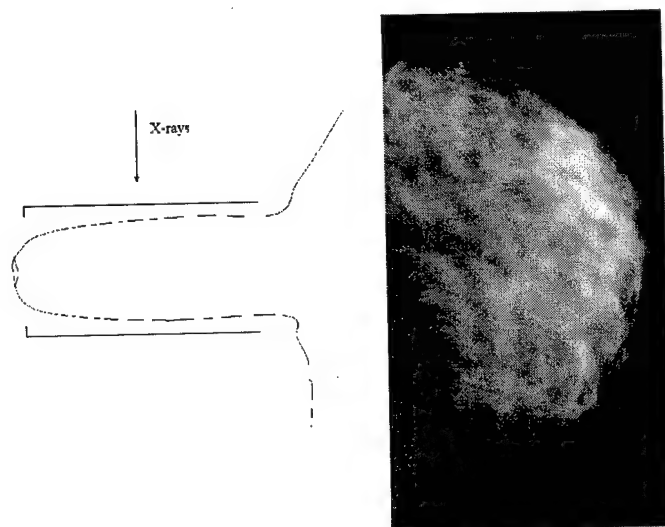
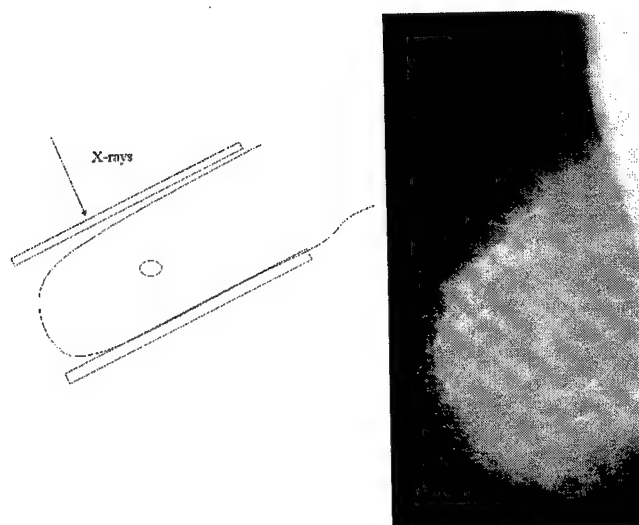Figure 2.3: Compression plain and sample CC mammogram view

Figure 2.4: Compression plain and sample mammogram for MLO view

Chapter 3

Patient Site model Construction and Update

3.1 Introduction

The site model is a dynamic mathematical and geometrical description of a scene under analysis. At a minimum, the site model contains the following parameters: objects, boundaries, object attributes, user input, and associated raw and processed data. The site model can vary in complexity ranging from detailed object description (building numbers) to simple boundary information. Pioneering work on the site model was performed by [13] in the analysis of aerial images for site monitoring and change detection for intelligence gathering purposes. In that research, the goals of site monitoring and change detection were accomplished through the support of three main model tasks. These tasks are called site model operations. In [79] the operations are defined as site model construction, image-to-site model registration, and site model update. Other research on the site model idea was performed by [74]. In [75], the site model operations are defined as site model acquisition, model-to-image registration and model extension. The pursuit of both of these research projects resulted in algorithms for automatic building detection [13], automatic and semi-automatic registration [79], [75], and fusion methodologies for combining user input with automatic processing results. Next, each of the site model operations will be further defined and discussed.

The first model operations is site model construction. Site model construction consists of deriving the site model parameters from the initial input images and user input. In [79], the construction process is as follows: (1) review two or more input images (overlapping views); (2) create a world coordinate system; (3) derive camera models for each image; (4) input camera focal length and principal point; (5) determine control points; (6) refine camera models for each image; (7) add objects and other annotations. [75] on the other hand, considers a lower level construction phase which includes (1) line segment extraction, (2) building detection, (3) multi-image epipolar matching, (4) multi-image triangulation, and (5) projective intensity mapping. These site model parameters which include detected line segments, buildings locations, camera models, and other control points are extracted using advanced and classical image processing techniques.

The next site model operation is image-to-site model registration. Image-to-site model registration is the process of putting a new incoming image (float image) into the same coordinate system as the site model (reference image). The registration process may be automatic or semi-automatic (user interaction). A general approach is to match, in some manner (via. criteria), selected site model parameters with newly extracted parameters in order to derive a transform that describes the recovering transformational geometry (transform) required for alignment. [79], [74], [78] describe several registration methods that they use with their site model. The result of this operation is an aligned image ready for change analysis.

The site model's ability to describe a scene over time is derived through the site model update procedure. Site model update allows for the addition of parameters (objects) of the site based on processing results of previous and current imaging conditions. With these operations, site change, such as a vehicle leaves a parking lot or lesion increase in size, can be detected and monitored efficiently. To maintain continuity, [79]'s notation for site model operation will be used throughout the rest of this report.

The site model idea can be extended to medical imaging analysis. In medical imaging, the radiologist often wants to perform similar types of applications to site monitor and change detection. For example, lesion detection and treatment monitoring. In these applications, a radiologist examines a temporal sequence (same view) of the same patient for change that could indicate cancer. When change is found further analysis is performed. For example, in mammogram screening, temporal sequences of the same patient are used to detect possible regions of interest.

Currently in medical imaging another type of model is used in various processing algorithms [52] called anatomical atlas (models). Although anatomical models are currently not used in change detection application, it is important

	Parameter	Size
1	Skin line	2XN
2	Raw image	MxM
3	Segmented image	MxM
4	Mask	MxM
5	Center Gavity	1x2
6	Eigenvalue	2x2
7	Eigenvector	2x2
8	Nipple location	1x2
9	Elongated Structures	MxM
10	Potenital Control points	nx2
11	Image histogram	1xMgl
12	change map	MxM
13	Internal objects	kx2xg
14	Control point Signatures	$\frac{N}{360}$xn
15	Quantification parameters	Kx3

Table 3.1: Site model parameters

to note the differences between the anatomical model and the site model. The main difference between the site and an atlas is the site model is specific to a particular scene (patient) where an anatomical model is more a textbook rendering of the scene that does not consider user input or individual variability. An example is an anatomical atlas of a MRI brain [57]. In this example, the synthetic brain MR image has the correct tissue percentages. This difference leads to a more refined name for the site model called the patient specific site model.

In this research, the site model is used to support registration and change detection to achieve the application goals of lesion detection and treatment monitoring in mammograms. The site model supports registration by providing a common frame (coordinate system) from which all other images in the sequence are registered. It also provides an efficient mechanism for combining manual site information (user label objects) and automatic information (detected boundaries and control points) in a useful manner to help facilitate the desired task. The rest of this chapter considers the specific contents of the model, the signal and image processing techniques used to construct the model parameters, and the site model update procedures.

3.2 Model Parameters

In this section, the site model components will be listed and their relevance discussed. Since the site model will be used to support sequence registration and change detection, it contains parameters used in the accomplishment of these tasks. Parameter order in the site model is arbitrary as the site model is interactive and parameters are used in a non-linear fashion. There are some parameters that depend on others, and naturally the dependent parameters would need to be calculated after the required information was available. The site model parameters included in this implementation are shown in Table 3.1. Next, the purpose of each parameter is discussed.

The first parameter is a $N \times 2$ vector containing the x, y coordinates of the breast skin line. The breast skin line parameter is used in initial registration as one of the multiple control objects and as the desired curve to be fit in nipple location estimation. The second parameter in the model is an $N \times 2$ vector containing the (x,y) coordinates of the largest objects, usually dense tissue, located within the breast tissue. These objects are used in conjunction with the skin line to perform multi-object registration. The third parameter is the $N \times 2$ contain the x, y locations of potential control points. These points are the cross points of horizontal and vertical structures (blood vessels and milk ducts) within the breast. The points are used to form the spatial-coordinate transform in the final registration phase. The fourth parameter is an image containing both horizontal and vertical structures. This image is used to generate point signatures for the determination of point correspondence between potential control points in reference (site) and float (incoming) image. The fifth parameter is the estimated nipple location and is stored in a 1×2 vector. The nipple location is used to localize point correspondence to a neighborhood window in the correspondence phase of final registration. The sixth parameter in the site model is the raw image histogram stored in a $1 \times MGL$ vector (MGL is the maximum intensity value in the image). The histogram will be used as the desired histogram in performing histogram specification between the incoming image and site. Histogram specification normalizes intensity ranges to that of the site model so object extraction is not biased by intensity differences. The eighth parameter is the image quantification model parameter estimates. These estimates are used
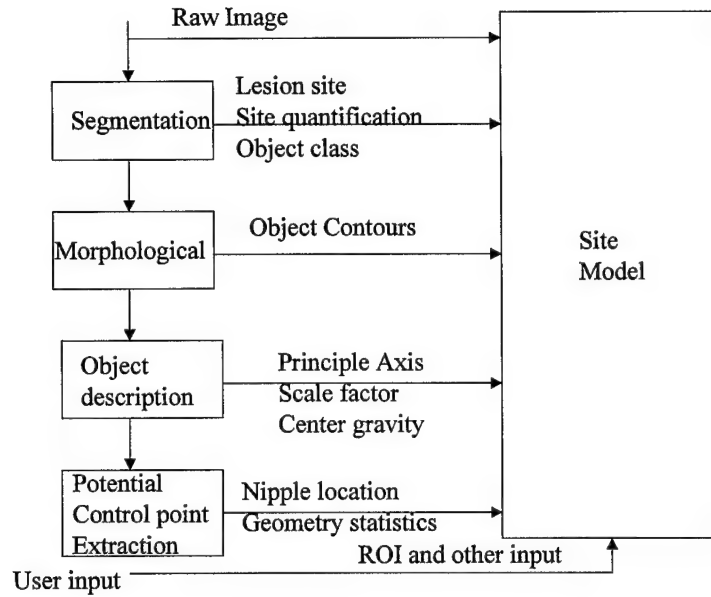
Figure 3.1: Site model construction flow

to initialize the segmentation of the incoming images so a uniform segmentation is achieved between the images. A copy of the raw image, raw segmented image and tissue mask are included and used in follow-on processing. Then, finally space is assigned for user specific input; such as the, number of classes in the scene, prominent landmark locations, change region of interest, and location of previous change. The number of classes in the scene is used to initialize the segmentation process. Prominent landmarks provide addition control points in final registration. The previous change location is used to exclude the change regions from further processing or focus in on specific regions for analysis.

The site model construction process is summarized in Figure 3.1. See Figure 3.2 for an example of a site model of a CC view mammogram. Next, the theory and algorithmic formation of each of the parameters will be discussed.

3.3 Model Construction

3.3.1 Segmentation

The segmentation algorithm used in this research is a statistical based algorithm that classifies each pixel as belonging to one of the K classes. The main premise of this algorithm is that the image's distribution can be represented by the gray level histogram of the image. The histogram of an image is defined as the number of times a pixel intensity falls within a pre-specified range as shown below.

$$p(u) \quad = \quad \frac{1}{N^2} \sum_{i=1}^{N^2} I(u, x_i) \tag{3.1}$$

$$I(u, x_i) \quad = \quad \begin{cases} 1, & u = x_i \\ 0, & u \neq x_i \end{cases} \tag{3.2}$$

where x is the intensity level of the pixel and I is an indicator function. Then it is assumed that the histogram can be mathematical modeled (or composed of) by a sum of K Gaussian distributions or mixture model where each individual Gaussian distribution identifies a class (tissue type). Finally, each pixel is assigned a class based on its membership probability. The algorithm is composed of two main components: quantification and segmentation. The quantification phase consists of estimating the parameters of the mixture model while the segmentation phase uses these estimates to determine pixel labels in a maximum likelihood sense.

Several studies of natural image statistics have yielded some stochastic image mixture models that best model the histogram of the X-ray mammographic images[19]. For this research we selected the standard finite normal mixture (SFNM) model as the histogram model. SFNM can be derived using the following relationships. First the
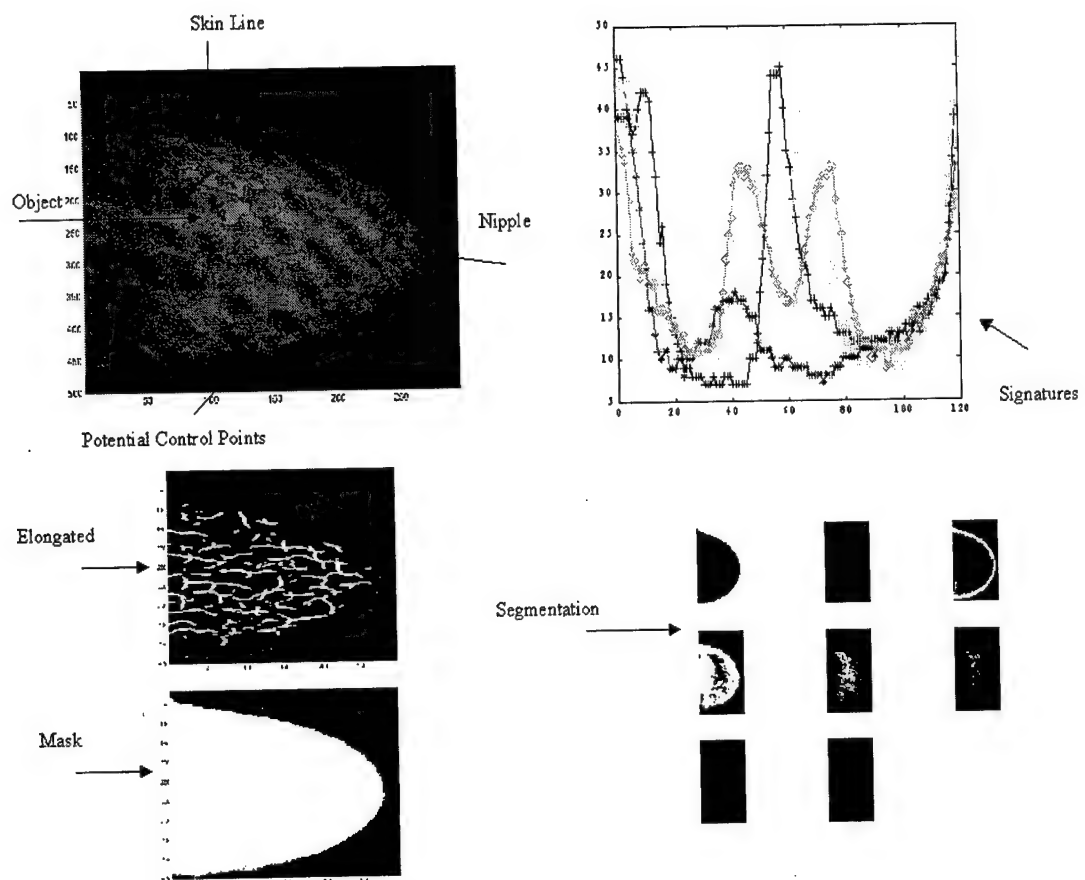
Figure 3.2: Example Site model

image is a $N \times N$ image where each pixel is assumed to be a random variable. The marginal distribution of the random variable (pixels) is shown below.

$$p(x) = \sum_{k=1}^{K} \pi_k \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right) \tag{3.3}$$

where x is the pixel (random variable), μ_k is the kth class mean, σ_k^2 is the kth class variance, and π_k is the distribution parameter. The SFNM is derived by randomly reordering the pixels with no regard to spatial information. This allows the pixels memberships to be treated as i.i.d. random variable. The joint distribution of the image is written as the product of each pixel's distribution as shown below.

$$P(X) = \prod_{i=1}^{N^2} \sum_{k=1}^{K} \pi_k \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x_i-\mu_k)^2}{2\sigma_k^2}\right) \tag{3.4}$$

The above equation represents the SFNM model which can be rewritten in the form of a likelihood function conditioned on θ, the free parameters vector.

$$P(X/\theta) = \prod_{i=1}^{N^2} \sum_{k=1}^{K} \pi_k g(x) \tag{3.5}$$

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{3.6}$$

$$\theta_k = \pi_k, \mu_k, \sigma_k^2 \tag{3.7}$$

In order to use this equation, the feature vector θ_k and K must be estimated. Since the components of θ are not treated like random variables, the estimation problem is formulated as a maximum likelihood (ML) estimation problem [76]. The main goal of ML estimation is to determine valves for θ and K that cause X to occur. Since the logarithm is monotonically increasing, maximizing the log-likelihood is equivalent to maximizing the likelihood function [76]. The ML estimate θ', is that valve of θ that maximizes the log-likelihood function. This estimate can be determined by differentiating the log-likelihood function $\log P(X/\theta)$ and setting it equal to zero (i.e. find the extreme point of the log likelihood function).

$$\frac{\delta \log P(X/\theta)}{\delta\theta}\bigg]_{\theta=\theta_{ml}} = 0 \tag{3.8}$$

Sometimes maximizing $\log P(X/\theta)$ is too complex to solve in a closed form solution. In cases like this, an iterative algorithm called the expectation-maximization algorithm (EM) can be used [25] to obtain the required ML estimates. The EM algorithm is designed to attack what is termed 'incomplete data problems' [25]. Incomplete data problems are defined as problems where part of the data for some reason is unobservable. Take, for instance, the true pixel labels L of an image as unobservable data and the pixels intensity Y as observable data. The relationship between observable and unobservable data is shown below

$$X = (Y, L) \tag{3.9}$$
$$X = T(L) \tag{3.10}$$

where X is the complete data and T is a nonreversible many-to-one transformation of L. If L could be observed directly then the complete information about the image would be known and no processing would be required. The EM algorithm is divided into a E step, where the likelihood unobservable data L is calculated through the observable data Y and the current parameter estimates, and a M step, where the unobservable likelihood function is maximized to yield new parameter estimates. In the SFNM formulation, the E step, for a assumed number of class K, this is formulated as a membership functions shown below

$$z_{jk}^{(m)} = \frac{\pi_k^{(m)} g(x_j/\mu_k^{(m)}, (\sigma_k^2)^{(m)})}{f(x_j/\theta^{(m)})} \tag{3.11}$$

where m is the current iteration number ranging from 0...... Then in the M step the updated parameters (μ, σ^2, π) are calculated by maximization of the likelihood with current estimates. The update equation are shown next.

$$\pi_k^{(m+1)} = \frac{1}{N} \sum_{j=1}^{N} z_{jk}^{(m)} \tag{3.12}$$

$$\mu_k^{(m+1)} = \frac{1}{N\pi_k^{m+1}} z_{jk}^{(m)} x_j$$

$$\sigma_k^{2(m+1)} = \frac{1}{N\pi_k^{m+1}} \sum_{j=1}^{N} z_{jk}^{(m)} (x_j - \mu_k^{(m+1)})^2$$

The EM iterates back and forth until a convergence criteria is reached (under regularity conditions) [25] . The convergence criteria is reached when the difference between $\pi_k^{(m)}$ and $\pi_k^{(m-1)}$ is smaller than some pre-determined value ϵ .

$$\left| \pi_k^{(m+1)} - \pi_k^{(m)} \right| < \epsilon \qquad (3.13)$$

A key factor in the use of the EM algorithm is obtaining a reasonable initialization of parameter estimates[25]. If initialization is not appropriate, then the algorithm could estimate into a local minima [25]. To combat this problem, the Adaptive Lloyd Max Histogram Quantization algorithm (ALMHQ) is used to determine the initial parameters estimates for the EM algorithm. [20]. The ALMHQ algorithm takes the image intensity histogram p and number of regions K as input then iteratively determines each of the K threshold values by trying to minimize the global distortion D with respect to the thresholds t and mean gray levels μ.

$$\frac{\delta D}{\delta t_k} = \frac{\delta D}{\delta \mu_k} = 0 \qquad (3.14)$$

$$D = \sum_{k=1}^{K} \int_{t_k}^{t_{k+1}} (\mu - \mu_k)^2 p(u) du \qquad (3.15)$$

After minimization of distortion, the update equations for μ can be derived as shown below.

$$\mu_k^m = 2t_k^m - \mu_{k-1}^m \qquad (3.16)$$

The σ^2 and π for each section are calculated once the optimal mean (μ) assignment has occurred. Iteration stops when the parameters no longer significantly change from iteration to iteration. These estimated values are used as the initial parameter estimates for the EM algorithm. The ALMHQ and EM assume that K is known however, except in controlled studies this is usually not true. The determination of K is termed a cluster validation problem[32] and can be solved using information criteria. The most commonly used information criteria is Akaike information Criteria (AIC). Appendix A describes this approaches along with some examples. Once the parameters have been estimated the quantification portion is complete. The results form the quantification phase are then used as input to the segmentation phase.

The segmentation portion consists of two main steps: maximum likelihood classification (MLC) which performs the initial segmentation, and contextual Bayesian relaxation labeling (CBRL) which performs the final segmentation [26]. The MLC can be used if we treat l_i^*, the true pixel label, as an independent non-random unknown constant. Then the label assignment is performed by maximizing the likelihood for each pixel in the image. The assignment of a pixel i into a class k is given by the following relationships

$$\Gamma(X/\mu_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(\frac{-(x_i - \mu_k)^2}{2\sigma_k^2} \right) \qquad (3.17)$$

$$l_i = \arg\{\max_k \Gamma(X/\mu_k, \sigma_k^2)\} \qquad (3.18)$$

where Γ is the likelihood function of pixel images for all pixels. The ML estimate of Γ for k would yield estimated k^{th}class label. This is realized by minimizing the log likelihood function given

$$d_{ik} = \log\left(\frac{1}{\sqrt{2\pi\sigma_k^2}} \right) + \frac{(x_i - \mu_k)^2}{2\sigma_k^2} \qquad (3.19)$$

where d_{ik} is defined as the Mahalanobis distance between the intensity of pixel i and mean of class k.

$$l_i = \arg\left\{ \min_k d_{ik} \right\} \qquad (3.20)$$

Thus, the label of the class mean that is closest to the pixel (in terms of Mahalanobis distance) is selected as the new pixel label.
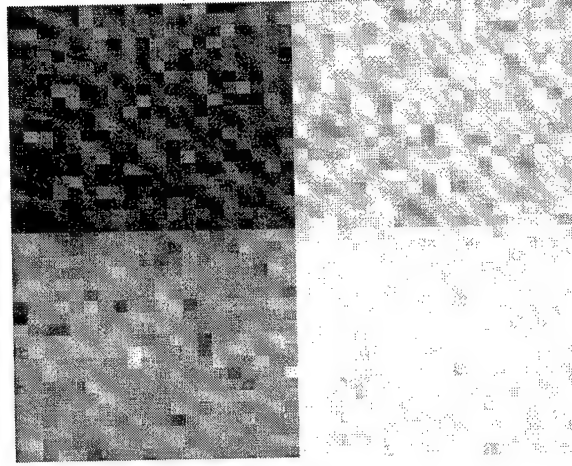
Figure 3.3: Raw four class phantom at 25db SNR

Relaxation labelling methods like CBRL perform efficient segmentation given initial pixel labels. This is accomplished by incorporating contextual information in the segmentation process. Context information is defined as the information relating a label (or class) to a pixel. The contextual information is considered by defining a neighborhood $b x b$ pixels around the pixel i. The CBRL derivation starts by defining δi the pixel neighborhood and $l_{\delta i}$ the labels of the neighborhood. $l_{\Delta i} = l_{j/\Delta i}$ $j = 1.....b^2$ $j = i$. Next, we can derive the neighborhood membership as

$$\pi_k = \frac{1}{b^2 - 1} \sum_{\Delta i} I\left(l_i = k, l_{j/\delta i}\right) \tag{3.21}$$

where I is the indictor function given by

$$I(x, u) = \begin{cases} 1, & x = u \\ 0. & x \neq u \end{cases} \tag{3.22}$$

π_k can also be interpreted as the conditional probability of l_i. The pdf of the gray level is given by

$$p(x_i / l_{\Delta i}) = \sum_{k=1}^{K} \pi_k p_k(x_i) \tag{3.23}$$

based on SFNM formulation. The segmentation is performed by minimizing the total classification error using the following relation.

$$l_i = \arg\left\{ \max\left(\sum \frac{I(l_{j/\delta i}, k)}{b^2 - 1} g(x/\theta_k) \right) \right\} \tag{3.24}$$

where $g(x/\theta_k)$ is the gaussian kernel.

3.3.2 Experimental Simulation

The quantification and segmentation algorithm was simulated with a phantom image and real mammograms. The phantom was a 40×40 image that contained four intensity values $(32, 42, 52, 62)$ each occupying 25% of the image. The image was then corrupted by AWGN that yielded a raw image with a SNR of 25dB as seen in Figure 3.3. The performance of the algorithm was evaluated by the analysis of the quantification and segmentation results. For quantification the true SFNM model parameters were compared to the estimated parameters. These results are depicted in Table 3.2.

From examination of the table the parameters estimates are within 0.5% error for μ and 7.0% error for π. Feeding the parameter estimates into the SFNM model and measuring the GRE between the phantom histogram and model shows that the distribution closely models the image. Finer estimates can be obtained, but the EM algorithm stop criteria must be deceased. In this current arrangement, the threshold is set to 5. By decreasing it to 1, the error

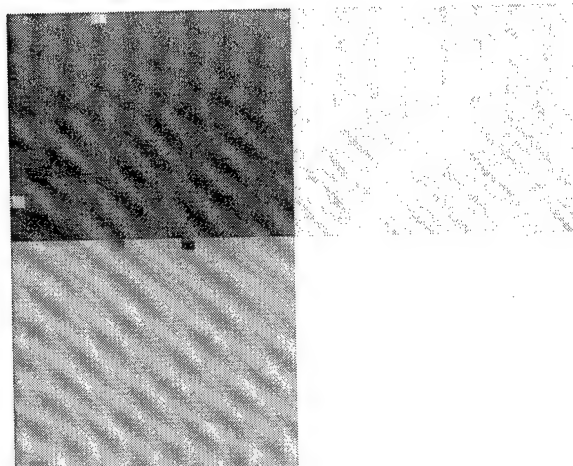| k	μ_k	$\hat{\mu}_k$	π_k	$\hat{\pi}_k$	$\hat{\sigma}_k^2$
1	32	31.82	.25	.242	6.9
2	42	41.79	.25	.2692	9.59
3	52	52.29	.25	.2460	6.7
4	62	62.08	.25	.2429	6.19

Table 3.2: SFNM parameters estimates for four class phantom.

Figure 3.4: Segmented version of four class phantom

percentages drops from 0.5% for μ to 0.3%. This error decease is also accompanied by an increase in processing time.

The results of segmentation of the four class phantom is shown in Figure 3.4. The performance of this portion of the algorithm was judged using the number of pixels in error and the amount of improvement in GRE between the processed and unprocessed images. In this example, the number of pixels in error drops drastically after processing from ___ to ____. This, in turn, improves down stream processing by removing unwanted intensity fluctuations in the image. This segmentation process is not without error. In several simulation runs, it appears that the error pixels are equally distributed across the image with most of the errors occurring between adjacent classes (i.e. pixels from class one are classified as pixel from class two). This appears to be attributed to the resolution of the quantification phase. This is similar to the resolution limitation of a FFT to resolve closely spaced frequencies [77]. If the quantification groups pixels into adjacent classes then the error feeds through into final segmentation.

The mammogram examined was 500×300 with 256 gray levels. From appendix A and [26], the number of classes for typical mammograms are found to be eight. Figure 3.5 shows a raw mammogram and Figure 3.6 shows a segmented version of the mammogram divided into individual classes. Because no ground true tissue map exist for real mammograms the performance will be compared to previous results obtained in [26]. Table 3.3 shows the estimates for the SFNM parameters for Figure 3.5.

These values roughly follow the results presented in [26]. Differences can be attributed to the imaging environment

	1	2	3	4	5	6	7	8
μ	27.39	32.89	62.84	105.17	132.24	159.53	181.45	203.55
σ^2	.46	1.9	294.69	162.75	82.38	81.00	76.04	52.06
π	0.0002	.353	.059	.052	.164	.116	.169	.087

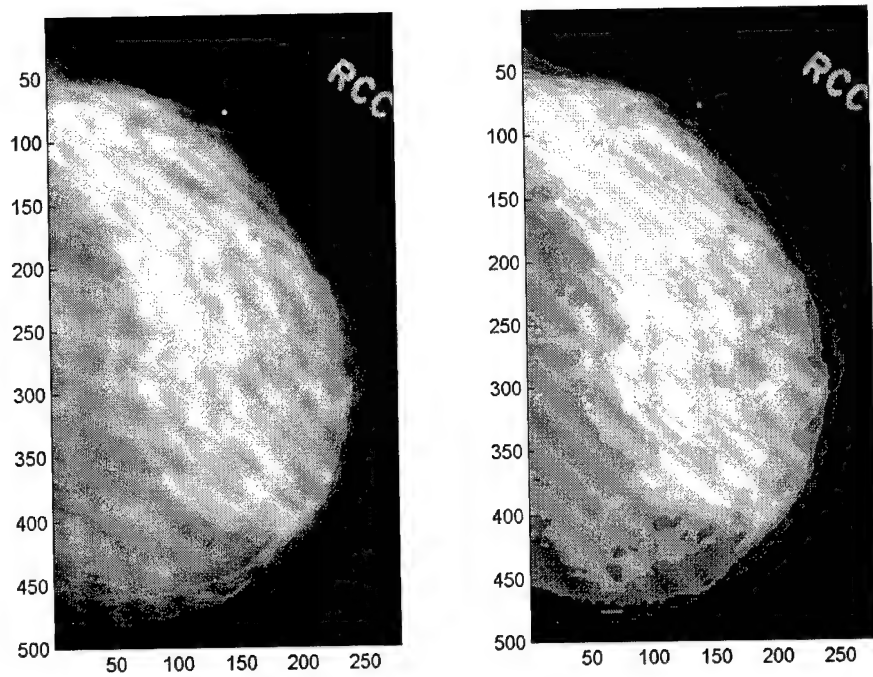Table 3.3: SFNM parameters estimates for mammogram with 8 classes.

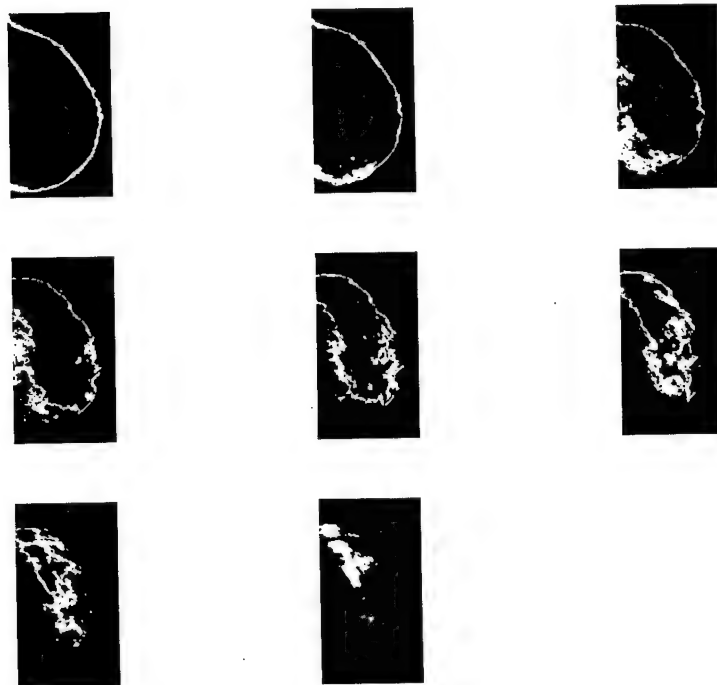Figure 3.5: Raw and segmented versions of a mammogram

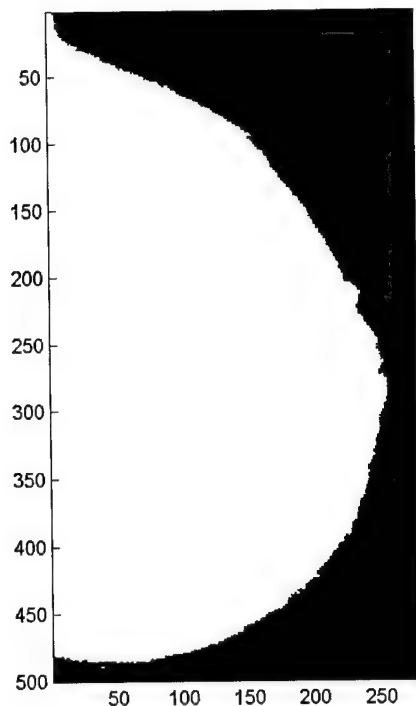Figure 3.6: Segmented classes from a mammogram

Figure 3.7: Mask image

(i.e. equipment used, signal strength, etc.).

3.3.3 Breast Tissue mask formation

The processing mask is formed by segmenting the raw image into two classes corresponding to tissue and non-tissue (background). Then for every pixel assigned to the tissue class the corresponding pixel location in binary image is set to one otherwise the pixel location is set to zero.

$$Mask(i,j) = \begin{cases} 1, & l_{i,j} = 1 \\ 0. & l_{i,j} = 0 \end{cases} \tag{3.25}$$

This mask image serves two purposes. The first purpose is to limit processing to only tissue regions of the image by multiplying non-tissue pixels by zero. This process increases processing speed and eliminates unwanted background effect in none tissue regions. The second purpose is to feed a morphological filter designed to extract the breast contour for use in further processing. Figure 3.7 shows some typical mammograms with the associated mask.

3.3.4 Contour Construction

The contour is constructed by passing the mask image through two morphological filters. Morphological filters are filters designed through a structuring element to perform different tasks. The structuring element is a $q \times q$ mask where $q \times q$ is smaller than the image size. The first filter is a dilation filter and it has the effect of thickening the object. The second filter is an erosion filter which has the opposite effect (i.e. thinning). The outline can then be formed by subtracting the dilated image by the eroded image yielding the object outline. A flow diagram of this process is shown in 3.8. Figure 3.9 shows some example extracted contours.

3.3.5 Object description

Initially point to point correspondence between images is unknown, but object to object correspondence is known. Using this object correspondence, an initial transform can be derived. Objects in the image include clustered dense
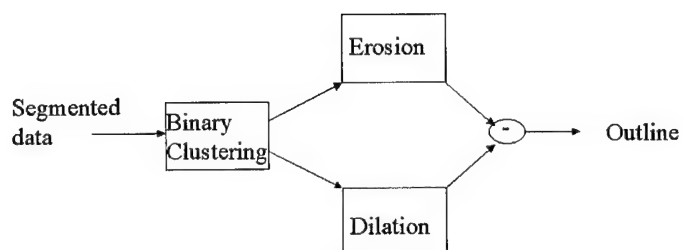
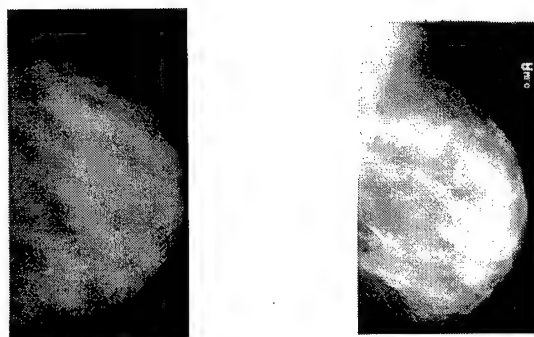Figure 3.8: Contour extraction process

Figure 3.9: Extracted mammogram contours.

tissue and the breast skin line. The first and second moment of the (x, y) coordinates are used to describe the object's geometry. The first moment is calculated using the following equation

$$R_g = \frac{1}{N}\sum_{i=1}^{N} r_i \tag{3.26}$$

where r_i is the (x, y) coordinate of a single point on the object of N samples and R_g is the center of gravity (first moment) of the object. The second moment is calculated using this relationship

$$C_r = \sum_{i=1}^{N} r_i r_i^t \tag{3.27}$$

where C_r is the covariance matrix (second moment) of the (x, y) points of the object. To further describe the object, the principle axes and dispersion along these axes is desired. The principle axes of a object is the axes about which the object's entia is minium. The dispersion along the axis is the spread of (x, y) values. The principle axis and dispersion have been shown to describe an object's orientation and scaling [53]. It has also been shown [50] that eigenvalue analysis [86] yields the principle axis and associated dispersions through the eigenvectors and eigenvalues of the covariance matrix of the object. So, the final description contains the center of gravity, principles axes and the dispersion along these axes.

3.3.6 Nipple point estimation

The nipple in most screening mammograms views lies on the extrema of the breast skin line. Several methods exist to determine the extrema point. In [88], the point is estimated by determining the point on the skin line that is farthest from the chest wall line. This method is suspectable to noise in chest wall estimation. Another more stable approach is by [7] which estimates the nipple location through least mean square error approximation of the skin line to a quadratic function. The skin line is obtained using intensity thresholding. The least mean square formulation is shown below.

$$f(x) = c_0 + c_1 x + c_2 x^2, \tag{3.28}$$

$$e = \sum_{i=1}^{n}(y_i - c_0 - c_1 x_i - c_2 x_i^2)^2, \tag{3.29}$$

$$\frac{\delta e}{\delta c_0} = 0, \quad \frac{\delta e}{\delta c_1} = 0, \quad \frac{\delta e}{\delta c_2} = 0,$$

where c's are weighting coefficients and n is the number of samples in the contour. The above derivatives yield the following system of equation where c_0, c_1, c_2 are the unknowns.

$$-2\sum_{i=1}^{n}(y_i - c_0 - c_1 x_i - c_2 x_i^2) = 0 \tag{3.30}$$

$$-2\sum_{i=1}^{n} x_i(y_i - c_0 - c_1 x_i - c_2 x_i^2) = 0$$

$$-4\sum_{i=1}^{n} x_i(y_i - c_0 - c_1 x_i - c_2 x_i^2) = 0$$

This approach is stable for breast skin lines that closely follow the quadratic function which MLO view images generally do not. In this research, the method by [7] is extended by the use of statistical segmentation to extract skin line, and a higher order polynomial as curve fitting function. The nipple estimation procedure is given by the following steps:

(1) Segment the raw image into classes.

(2) Group those classes into two classes of breast tissue and background forming a binary image.

(3) Extract the skin line using morphological filtering.

(4) Using N contour points $f(x_i)$ of skin line, curve fit a nth order polynomial using least squares. The formulation is as follows:

$$f(x) = c_0 + c_1 x + c_2 x^2 + \ldots\ldots c_n x^n \tag{3.31}$$

$$e = \sum_{i=1}^{n}\left(y_i - \left(c_0 + \sum_{l=1}^{n} c_l x_i^l\right)\right)^2$$

Method	x	y
GOOD	289	279
LEHIGH	275	294
WOODS	287	277

Table 3.4: Estimated nipple locations for a CC contour the methods.

Method	x	y
GOOD	345	278
LEHIGH	238	236
WOODS	367	274

Table 3.5: Estimated nipple locations on a MLO contour for the three methods.

This leads to a $n+1$ system of equation to be solved for the weighting coefficients c .

(5) Find the critical points of $f(x)$ using the following

$$\frac{df(x)}{dx} = 0 \qquad (3.32)$$

then solve of x.

A nth order polynomial results in $n-1$ roots. So, to reduce the number of roots to a manageable number complex roots, zero roots, and roots outside the breast tissue were dropped from analysis. The x yielding the largest $f(x)$ is selected as the skin line extrema or nipple location.

3.3.7 Simulation Experiments

The performance of this algorithm was tested through comparison with the results obtained by [6] and [88]. The skin line contours were extracted using the procedure describe in above section. The algorithms were run on several CC and MLO view mammograms. Table 3.4 shows the x, y location for a representative CC mammogram using the three methods.

Table 3.5 show the x, y location for a representative MLO

In the CC image, our method obtains a nipple estimate closest to the visually selected nipple, but in the MLO image the [88] method selects the best nipple. Our method selects the bottom of the nipple in this case. On average, our method out performs both [6] and [88] because of the low order polynomial used for curve fitting and contour extraction noise. Table 3.6 shows the MSE between a contour and various order polynomials functions for CC and MLO mammograms.

From this we see the higher order functions obtains a lower MSE especially on MLO contour which are not generally quadratic. Thus, with higher order polynomials a more robust nipple estimation is achieved. To further highlight the need for higher order polynomials, Figure 3.10 shows the nipple locations given various order polynomials. The proposed algorithm results were evaluated by radiologists and were found to be accurate in 95 % of the cases. Although some cases estimated the top or bottom of the nipple, the 5 % error can be attributed to contour extraction error. In these cases, the contour was not very smooth causing many local extrema points. This problem could be addressed using a smoothing filter on the contour before nipple detection.

CC			MLO	
Order	MSE		Order	MSE
2	415.9		2	3640
5	162.8		5	1419
10	113.4		10	1381
20	415.9		20	1129

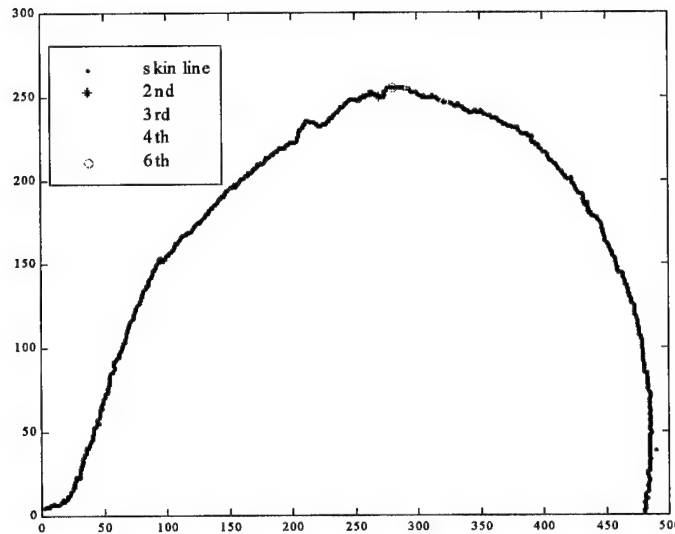Table 3.6: MSE between the contour and n^{th}-order polynomial for CC and MLO views

Figure 3.10: Estimated nipple locations for 2,3, 4, and 6th order.

3.4 Site Model Update

In [75] and [13] site model update (or extension) is the process of finding and modeling un-modeled buildings (objects) and adding them to the site database. This is possible because the image-to-site model registration provides the correspondence (overall alignment and camera angle) necessary to compare regions. Once registration is completed, the newly aligned images are then processed looking for set model parameters. These new parameters are compared to existing parameters looking for differences. The differences in parameters are new locations which are then added to the database yielding a composite view of the scene.

In this research, the use of the site model differs from that of [75] and [79] because the site model is used as a reference model with a variable parameter (change map) not a variable model where every parameter could be updated. Site model update, for this application, identifies changes found in new images (registered) and adds them into the site model parameter change map while leaving the other site model parameters untouched. The untouched parameters represent the characteristic of the reference image, and by definition of reference should not be altered. So, overtime this database will contain the reference image information and changes that have taken place over the sequence. This formulation of the site model meets the main objects stated previously which are to provide a common registration frame and highlight the change region for possible exclusion from further processing. Next, the update processes will be explained in more detail.

The site model update process is conducted by modifying the change map (M) parameter with the newly detected change. The change map parameter is an image the same dimension as the scene image where each pixel $M(i,j)$ is initialized to zero to start. Then, each time a pixel $M(i,j)$ is identified as being changed the value of $M(i,j)$ is incremented. Figure 3.11 shows an example change map for a growing object. From Figure 3.11 we see that the object has grown through four images of the sequence. This map could then be used to quantify the change by calculating the size, shape, and rate of change for the object through the sequence.

1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	2	2	2	2	1
1	1	2	2	2	2	2	2	1	1
1	1	2	3	3	3	3	2	1	1
1	1	2	3	4	4	3	2	1	1
1	1	2	3	4	4	3	2	1	1
1	1	2	3	3	3	3	2	1	1
1	1	2	2	2	2	2	2	1	1
1	1	2	2	2	2	2	2	2	1
1	1	1	1	1	1	1	1	1	1

Figure 3.11: Change map for a 4 image sequence

Chapter 4

Site Model Supported Image Registration

4.1 Introduction

The registration process is supported by the concept of a site model and site model operations. The site model is a mathematical representation of a scene under analysis. A basic site model contains a geometric description of an scenes objects (area, size, and other attributes), raw data, and simple user input (previous tumor locations). The environment interacts with the site model through the site model operations: construction, image-to-site registration and model parameter update. The site model is constructed by thoroughly processing the first image in the sequence to obtain the parameters. The site model supports registration in three main ways. First, the site model forms the reference frame (reference image) for all subsequent images, thus allowing all of the images in the sequence to be alignment to a common coordinate system. Second, the model stores registration parameters like object contours, control points, and user identified regions. This effectively integrates both manual and automatic control objects in a single place. Third, the model stores previously detected change, this enables the current registration process to exclude the previously detected changed portion from the current analysis which improves algorithm robustness. This chapter mostly considers the development of the image-to-site model operation starting with registration theory.

Image registration is the process of overlaying two images with the motivation of transforming one of the images, usually called the float image ($I2$), into the same coordinate system as the other image called the reference image ($I1$). The process consists of two steps. First, perform a spatial-coordinate transform or mapping function (f) which is used to determine the corresponding coordinate in the new image as shown below.

$$(x', y') = f(x, y) \tag{4.1}$$

In more complex mappings, f can be broken into f_x and f_y corresponding to the x-component and y-components respectively. Typically, (x', y') will not map to an integer grid point on the new image so, some interpolation is need to find the correct (x', y'). The second step of registration is the intensity transform (g), which is used to assign an intensity value to the pixel location (x', y'). Interpolation of the gray levels may also be required to obtain the intensity of point (x', y'). The mathematical expression for registration is given next.

$$I_2(x', y') = g(I_1(f(x, y))) \tag{4.2}$$

Some registration application do not require an intensity transform (i.e. intensity mapping table) such as single modality registration with similar gray level distributions, but multi-modality applications require a more complex transform that accounts for gray level differences between the two modalities.

The key problem in image registration is the determination of the spatial-coordinate transform. The most common types of transforms are rigid (distance between points in the image are preserved under a transform); affine (straight lines and parallelism are preserved between images); projective (straight lines are preserved); and curved (straight line on the original image maps to a curve on the new image). The rigid transform is characterized by a rotation, translation, and scaling which is realized by the following relationship:

$$F = AX + T \tag{4.3}$$

where A is the rotation matrix and T is the translation matrix. This equation can be rewritten as the following

$$f(x,y) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \tag{4.4}$$

$$a_{11} = a_{22} = \cos(\theta), a_{21} = a_{12} = \sin(\theta), a_{13} = t_x, a_{23} = t_y.$$

The affine transform is more flexible because the a values from the above equation are not restricted to take on only *sin* and *cos* values. The only constraint is A must be real valued. Projective transforms are realized in a similar manner

$$f(x,y) = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ l \end{bmatrix} \tag{4.5}$$

where w is the extra homogeneous coordinate. Finally, the curved transform is modeled by a nth order polynomial as shown below.

$$f(x,y) = a_{00} + a_{10}x + a_{01}y + ...$$

In complex mappings, each axis (x-axis, y-axis) has its own polynomial defined as $f_x(x,y)$, $f_y(x,y)$. These polynomials can model several types of transforms. In this research, we focus on the rigid, affine, and polynomial based registration methods to register the sequence of mammograms of the same patient.

Image-to-site model registration is performed by a multi-step algorithm consisting of an initial and final phase. The initial phase registers the images using the principle axis of the skin line in conjunction with segmented internal objects to form a multi-object global rigid spatial-coordinate transform followed by a simple look up table for the intensity transform. The final registration phase consists of a global thin-plate spline transform derived from the control points of the interior breast tissue. The intensity transform in this step is also a look-up table. Next each phase is described in detail, followed by simulation, results, and discussion.

4.1.1 Initial Registration

The main goal of initial registration is to correct for large mis-alignments between images in a sequence. The mis-alignments come from differences in breast placement upon examination, image acquisition process, and film size differences. Although the breast is generally considered a non-rigid object [84], a rigid approach is used as the basis of this phase. This approach is justified by the fact that the distortions, the initial phase is trying to correct, are more or less rigid in nature. In addition, it can be applied without complex knowledge of the input data (i.e. control point correspondence). An example change that is consider my this phase is film size differences. This occurs when different film sizes are used in the acquisition. This type of problem is handled by increasing or decreasing the image by a global scale factor which is addressed by a rigid transform (scaling). The initial registration is performed by a multi-object principle axis registration (PAR) algorithm. The objects include the breast skin line and other extracted internal objects (i.e. clustered dense tissue). The algorithm proceeds as follows: (1) Extract the contours (skin-line and internal objects) from both images. The contours and objects from the reference image are stored in the site model. (2) Use PAR to obtain the transforms for each object. To insure similar objects are extracted from both images, the incoming images are histogram specified to match the reference image (site). (3) Transform each pixel of the image using the transform that is closet in terms of Euclidean distance. This type of transform is called a local rigid transform. The complete process can be summarized into three main steps which are preprocessing, spatial-coordinate transform, and intensity mapping. Figure 4.1 shows a flow chart of the initial registration process. Next, each of these phases are explained.

4.1.2 Preprocess

In this phase, the objects used in initial registration are determined. An object is defined, as a cluster of the same tissue type in the image. Tissue types are identified with statistical segmentation which assigns a label (tissue type) to each pixel of the image [19] [26]. Clusters are identified by using class based region growing where the joining criteria is the pixels class membership. In order to perform registration, some level of correspondence must be established between the images. Visual inspection of extracted objects is used to determine object correspondence. An important step in this process is the identification of similar objects. This problem can become complex when the two images have different pixel intensity ranges. This causes the segmentation algorithm to produce different pixel class assignments resulting in different looking objects. To combat this problem, histogram specification is performed on the incoming image in order to match the site image. In histogram specification, the goal is to adjust
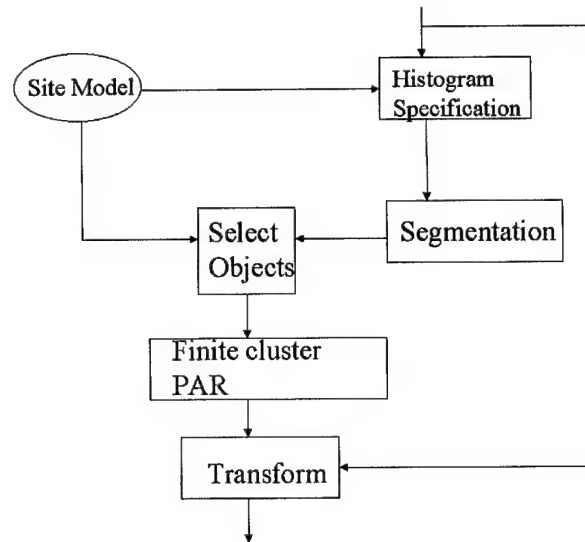
Figure 4.1: Process flow for initial registration phase

the intensities of an image so that the image's histogram matches a desired shape namely the histogram of the site image [85]. The process consists of three steps:

1. Equalize the input image histogram via histogram equalization [85].

In histogram equalization, the raw image intensity values are adjusted to produce a uniform histogram. Consider the pixels x in the image to be random variables with a probability density distribution of $p_x(x)$ and a cumulative distribution of $F_x = P[x <= x]$. Then an associated uniformly distributed random variable would be $y = \int_0^x p_x(x)dx$. In the digital domain, the integral is replaced by a sum which results in the follow equation.

$y = \sum_{i=0} p_x(i)$ where y is the new pixel value resulting from the transform $y = T(x)$.

2. Equalize the desired histogram (histogram of site image).

3. Determine the new gray level by matching the pixel value in the equalized image y with the gray level required to make the transform equate to y. $y = G(z)$ $z = G^{-1}(y)$ where z is the new intensity level and G is the transform

Now the histogram specified image is then segmented resulting in more similar looking class assignments.

4.1.3 Simulation Experiments

Next, an object extraction example is consider using the sequence shown in Figure 4.2. This sequence is composed of mammograms of the same patient, acquired at different times. Figure 4.3 shows the class assignment for Figure 4.2. From this figure we see the segmentation did not yield uniform pixel membership across the sequence. Thus, object selection becomes subjective. This fact is highlighted by examining the histograms of the images as shown in Figure 4.4. To correct this problem, the incoming histogram is specified to better match the site image. This is shown in Figure 4.4. This results in a uniform segmentation across the sequence as seen in Figure 4.5. Region growing is then applied to both images to create the objects. Objects from Figure 4.2 are shown in Figure 4.6 and 4.7. The objects are then used in the calculation of the spatial transformation.

4.1.4 Spatial transformation

The transform is calculated by using principle axis methodology[50]. The principle axis method is based on determining and manipulating the principle axes of an object in an image. The principle axis of an object is the axis about which the moments of inertia of the object are minimum. In this method, the objects are registered by matching the principle axes. This approach only works with objects that only vary in rotation and scaling. The rotation factor is represented by the eigenvectors of the data scatter matrix and the scaling factor is address by the ratio of associated eigenvalues of the scatter matrix. Translation is handled by collocating both objects at the origin. The algorithm is as follows: (1) obtain the associated coordinates of the object of interest in both images. (2) Determine the center of gravity object using the following equation.
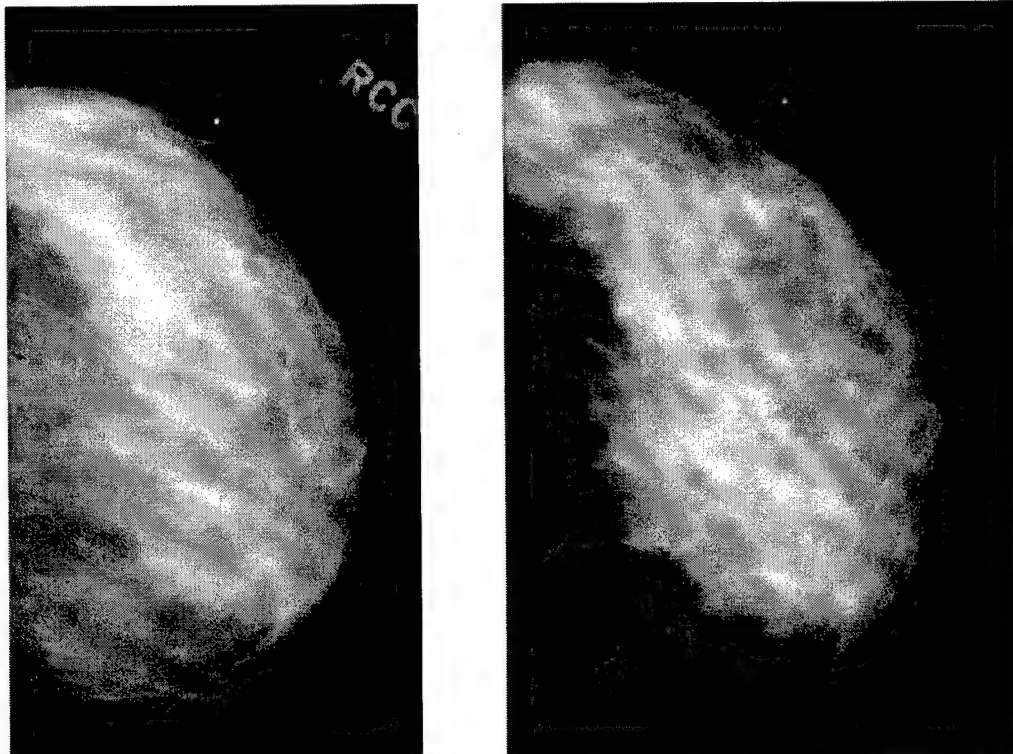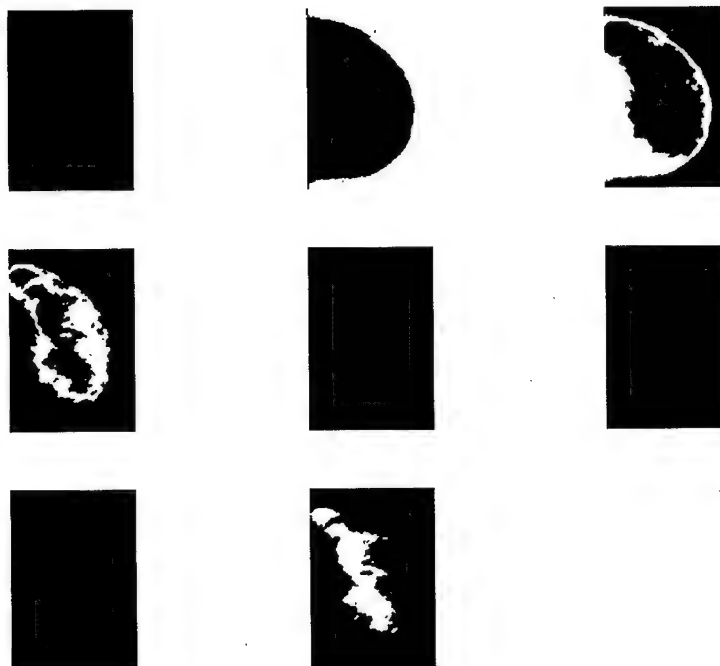
Figure 4.2: Squence of mammograms
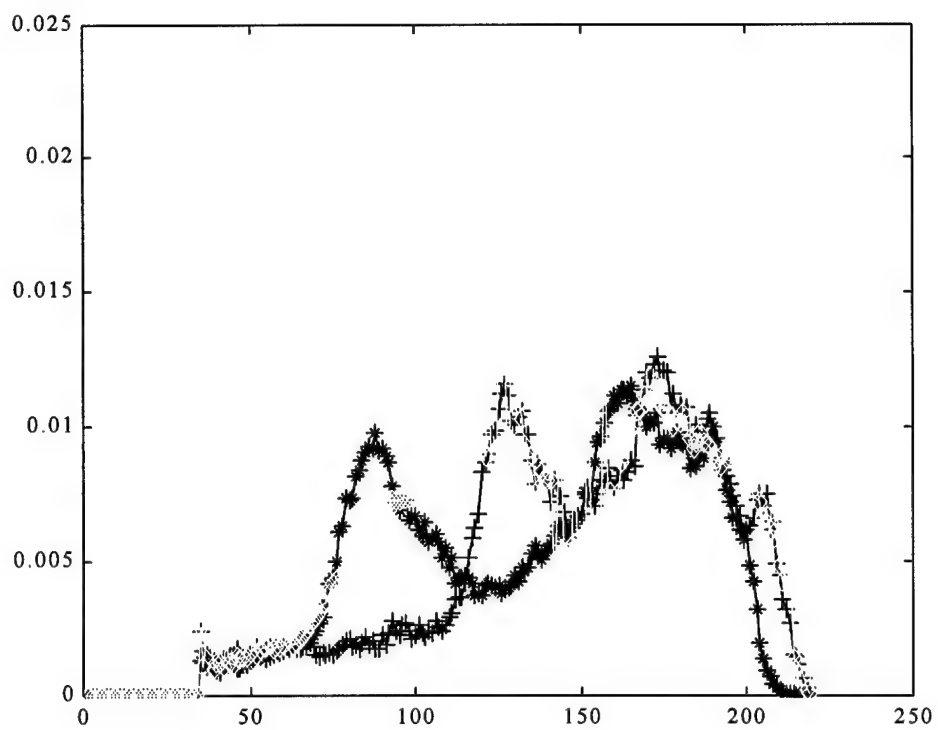

Figure 4.3: Class assignment for raw squence
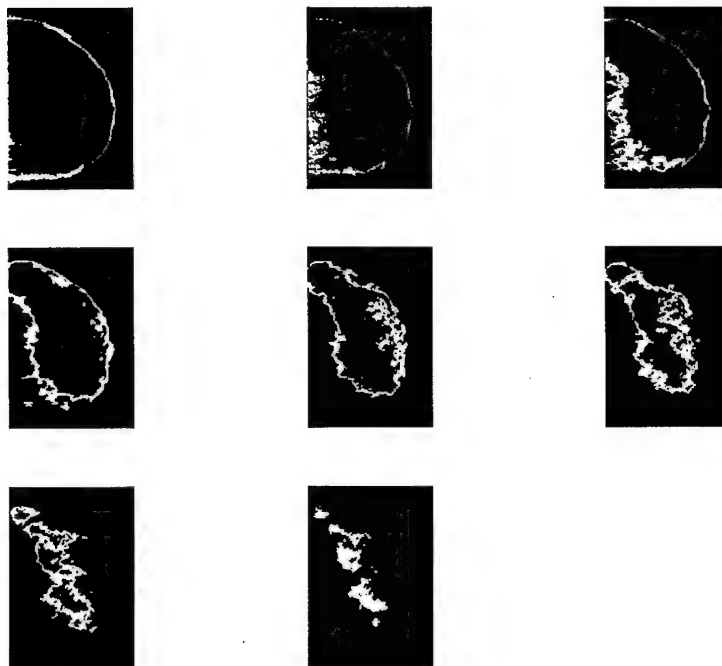
Figure 4.4: Plots of histograms
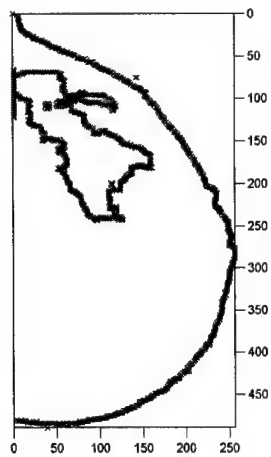

Figure 4.5: Class assignment for specificied image

Figure 4.6: Selected object in the site image.

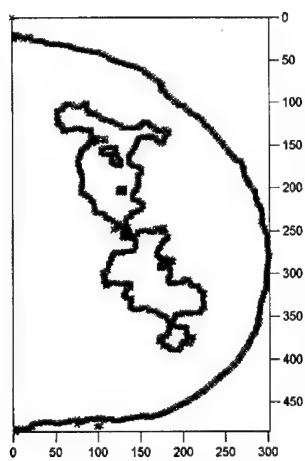Figure 4.7: Selected object in the float image.

Set	UnRegistered	Registered
1	48.015	37.391
2	45.354	39.613

Table 4.1: Mse between registered and unregistered contours

$$r_{cg} = \frac{1}{N} \sum_{i=1}^{N} r_i \qquad (4.6)$$

r_i represents a point (x, y) and N is the total number of points in the object. (3) Translate the objects so the center of gravity of each object is the origin $(0,0)$ given by q_i

$$q_i = r_i - r_{cg} \qquad (4.7)$$

(4) Calculate the scatter (covariance) matrix of the translated data points q_i's.

$$M = \frac{1}{N} \sum_{i=1}^{N} (q_i)^T q_i \qquad (4.8)$$

(5) Search for the transformation matrix that diagonalizes M. The transform matrix will be composed of the eigenvectors of M (principle axis). This can be realized by performing singular valve decomposition (SVD) of M

$$\Lambda = V^T M V. \qquad (4.9)$$

where Λ is a diagonal matrix containing eigenvalues and V contains the associated eigenvectors. (6) Determine the scaling matrix by forming a ratio between the axis dispersion (eigenvalues) of each image.

$$\Phi_f S^2 = \Phi_r \qquad (4.10)$$

where Φ is the diagonal matrix containing the eigenvalues and S^2 is a diagonal matrix contain scale factors for each axis. (7) Form the final transform which is a combination of rotation and scaling which is given below.

$$U = V_f^T S V_r \qquad (4.11)$$

4.1.5 Simulation experiments

This portion of the system was simulated using the skin line contours of the breast as objects. The derived transform was then applied to the contour points of the float image to obtain a transformed contour. The performance is measured by the MSE between the contours as shown in Table 4.1. Figure 4.8 shows two examples with raw unregistered contours with the associated warped contour. From this table and figure it is apparent that after registration the contours are spatially closer together. The difference between the mse for registered and unregistered is only be about 22%. This is attributed to the end effects where contour points at the beginning and end of the contour create large amounts of matching error. Reducing focus to only consider the central portion of the contour would significantly increase the difference between registered and unregistered mse.

4.1.6 Combination of Spatial Transforms

Assume that multiple corresponding objects can be extracted from the image pair, and from these objects control points could be determined using either contours, surfaces, or object description. In registration, these control points are used to determine a spatial-coordinate transform T that maps pixel in one image to pixel in another. The general expression is shown below

$$x_i^{'} = T(x_i) \qquad (4.12)$$

where $x_i^{'}$ is the transformed pixel and x_i is the pixel to be transformed. Three combination approaches have been investigated during the course of this research. Approach one, is a standard approach that considers each of the object pairs as separate registration problems yielding a transform for each object pair. Then a pixel is transformed by a particular transform via some metric Θ (i.e. pixel to contour distance).

$$\begin{aligned} x_i^{'} &= T_k(x_i) \\ k &= \Theta(x_i, T_l()) \quad l = 1, ... K \end{aligned} \qquad (4.13)$$
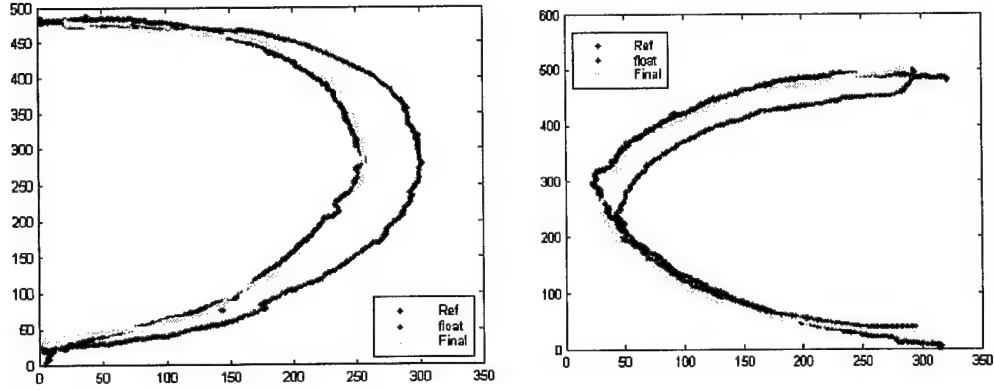
Figure 4.8: Unregistered and registered breast contours.

where k is the transform index ranging from 1 to K number of transforms calculated. This type of transform is called a local rigid/non-rigid transform because pixels are transformed based on transforms local to the pixel [73]. The second approach assumes that each of the T_k describes the same transformation. Then the final transform is obtained by average The signal model is given below

$$t_i = f_i + w$$

where f is the transform and w is the noise.

Signal averaging is routinely used to improve the signal to noise ratio of signals that are corrupted by noise and can be measured repeatedly [77]. In our case we average the transforms created from all of the objects under analysis to obtain a master transform (T) which is applied to the complete image.

$$T = \frac{1}{K} \sum_{i=1}^{K} t_i$$

where t_i represents a sample transform and K equals the total number of transforms in the image. This method leads to a global rigid/non-rigid transform since each pixel is transformed by the same matrix.

The third approach, considers the control points as belonging to one of K clusters each with its own mean and variance. Using the mean and variance each cluster can be modeled as a normal distribution. Now, instead of the pixel x_i only being influenced by a single transform it is influenced by a multiple transforms specifically K . The standard transform equation shown above is modified as follows.

$$x_i^{'} = \sum_{k=1}^{K} \alpha_{ik} T_k(x_i)$$

where α_{ik} is the weighting factor for the ith pixel for the kth transform. This formulation reduces back to the standard transform equation when $\alpha_{ik} = 1$; $\alpha_{lk} = 0$; $l \neq i$; Thus each $x_i^{'}$ in this formulation is the weighted sum of K transforms. The weight function could take on several forms such as distance, average, or probability membership. Given that the control points are localized to clusters described by their mean and variance, all of the control point clusters could be made to define a finite normal mixture model as shown below
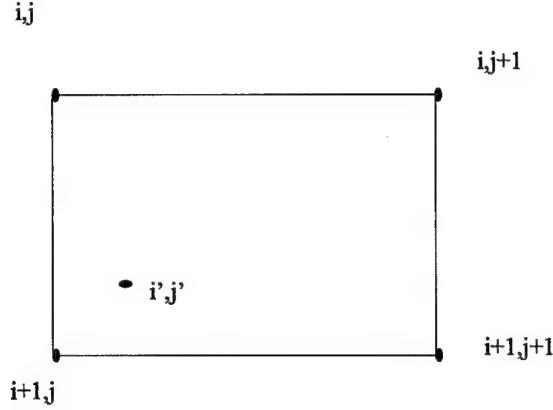
i,j

i,j+1

• i',j'

i+1,j

i+1,j+1

Figure 4.9: Four pixel grid with point (i',j') that falls between the points

$$f(x) = \sum_{k=1}^{K} \frac{1}{K} g(x/\mu_k, C_k)$$

where g is a gaussian kernel and μ_k and C_k are the class mean and covariance respectively. The mixture model sets the framework for using pixel membership as a weighting criteria. Membership in this context is defined as which transform is used to transform a pixel. This model has been used in image segmentation to determine pixel class [19] [28] [54]. Similar to [19] [28] [54] the posterior probability is used as a measure of each pixels probability membership. The statistical membership of a pixel with respect to a control point cluster can be defined as

$$\alpha_{ik} = P(T_k/x_i) = \frac{g(x_i/\mu_k, C_k)}{\sum_{l=1}^{K} g(x_i/\mu_l, C_l)}.$$

Thus each pixel in the float image can now be transformed using a membership weighted transform. The gray levels of each pixel are assigned using a straight look up table. The procedure is the following: (1) transform the pixels located at a point (x, y) in the reference image (R_I) to a point (u, v) in the float image (R_f) using the selected transform (T).

$$(u, v) = T(x, y)$$

Determine the intensity at point (u, v). Since points (u, v) are generally not integer values (i.e. fall on a grid point), interpolation is required to select the intensity. Figure 4.9 highlights an example which requires interpolation. Several interpolation method exist, but for this research Nearest Neighbor interpolation is used. This method assigns the new value (u, v) from the closets grid point surrounding it. This leads to the following relationship.

$$w(x, y) = R_f(T(x, y))$$

4.1.7 Simulation Experiments

The implementation of the following methods are discuss through some examples. Figure 4.10 shows the original image pair under consideration. The image pair was created by the addition of a Gaussian filtered block and rigidly rotating the complete image by $10°$. This is a small rotation, but should highlight the effect of the local and global multiple object transforms on the image. Figure 4.11 and 4.12 shows the resulting image pairs after transformation by the local rigid and global rigid transform respectively. From examination of Figure 4.11 it is apparent that discontinuity resulted from the transform as seen on the left hand side of the right image in Figure 4.11. These discontinuity can be attributed to differences in transform used on adjacent pixels. The global registration pair, on the other hand, has a smooth look because of the use of a single transform. So, no more cases of adjacent pixels
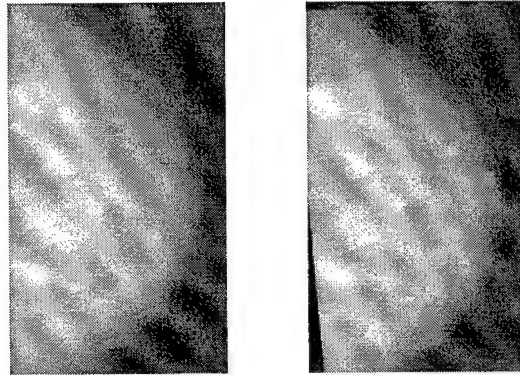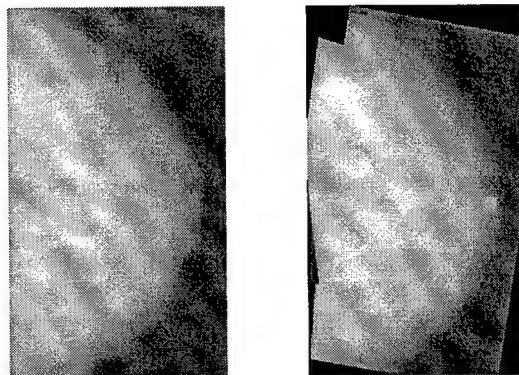
Figure 4.10: Original image pair

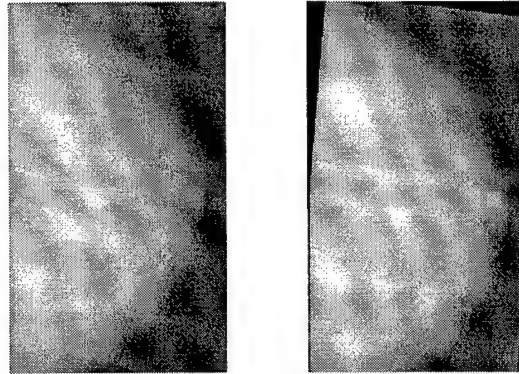Figure 4.11: Image pair transformed using local rigid with three objects

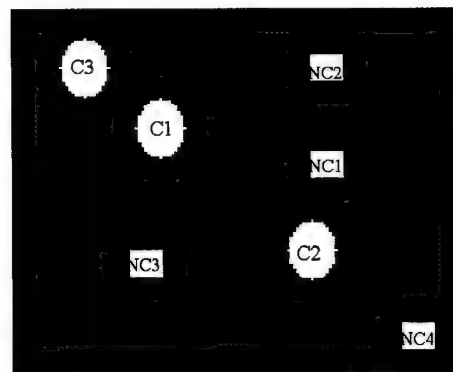Figure 4.12: Image pair produced with the global rigid

Figure 4.13: Phantom image used in finite normal mixture registration

Object	Configuration		
	1	2	3
C 1	5	7	20
C 2	100	4	5
C 3	20	25	10
NC 1	10	13	4
NC 2	20	6	24
NC 3	5	20	11
NC 4	5	20	11

Table 4.2: Angle rotations for each object in phantom registration image

Number of Objects	Config. 1	Config. 2	Config. 3
0	1616	1598	1785
1	960	930	468
2	758	842	410
3	279	504	310

Table 4.3: Mse results for each configuration

being transformed by different transforms. To simulate the finite mixture registration method, we considered a 150 x 150 phantom image containing three control objects and four non-control objects as seen in Figure 4.13. The control objects are ellipse while the non-control objects are squares 10×10. The key thing about the control objects is that only object correspondence is known not point correspondence. Each of the control and non-control objects are rotated and translated by different amounts. This simulates a non-linear deformation (non-rigid) between image sets, and serves to test the combination ability of this registration method. The objects rotation angles are given in Table 4.2.

Three configurations of rotation angles are considered. These configurations are chosen arbitrary to show the robustness of the proposed algorithm. In each configuration the image is registered using one, two, or three transforms. The performance is measured in mean square error (mse) between the reference and warped image where a lower mse is seen as better performance. Table 4.3 shows the mse for each configuration. From the table it is apparent that registration by one transform on average reduces the mse by 50%. The mse is deceased another 10% with the addition of another transform. With the addition of the last transform, significant improvement in mse is achieved. The mse is reduced by approximately 75%. Figure 4.14 shows an example of the reference and warped image using all three transforms. These results show the benefit of using multiple transforms where possible.

4.1.8 Final Registration

The goal of this section is to fine tune the alignment achieved in the initial phase by considering the breast as a non-rigid body. This allows for the consideration of the deformation between the image and site model. Deformations are caused by positioning differences subject weight gain, natural growth, and nonuniform compression during examination. To handle these deformations, more complex transforms are required. In [68], the polynomial based transform were shown to be able to handle non-rigid deformation of kidneys so they are selected in this study to model the deformations of the breast. Various types of polynomial transforms exist such as linear, quadratic, and cubic [68]. In this research, a thin-plate spline polynomial will be used as the mapping function [5].

The key requirement for use of polynomial based transform is the existence of control points. In some environments control points are easily obtained (brain images), but in mammograms this task is very difficult because of lack of anatomical landmarks between images. In this research, the cross points between vertical and horizontal elongated structures are used as potential control points. These elongated structures represent blood vessels and milk ducts. To use these points, one must assume they are time and shift invariant for the most part. These points will be defined as potential control points. Then the potential control points are matched to produce the final control points which are then used to calculated the thin-plate spline polynomial transform. The fine registration process concludes with the transformation of the complete image pixel by pixel.

Similar to the initial registration , final registration can be divided into several parts. They are preprocessing, point correspondence, spatial coordinate transform, and intensity mapping. Figure 4.15 shows the complete
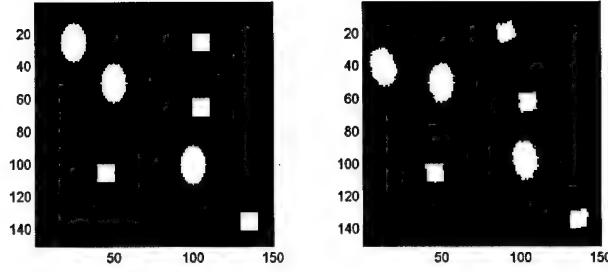
Figure 4.14: Reference and warped image from multi-object registration

algorithm flow. Next, each part will be discussed in detail.

4.1.9 Preprocessing

In this part, the potential control points are extracted from the image. This is achieved by detect the elongated structures in the image using modified monotony operators to highlight both horizontal and vertical structures in the image[7]. The monotony operators are defined by two overlapping rectangular neighborhoods, one small and one large, centered around a pixel (i, j). Figure 4.16 shows an example of both the vertical and horizontal operators in a image. The operators work as follows: the pixel at (i, j) is labeled one if the number of pixels in the large neighborhood that are larger than g_{max}, exceeds a threshold τ. Otherwise, the operator assigns a zero to the pixel (i, j). g_{max} is defined as the maximum gray level in the small neighborhood surrounding the pixel (i, j). The vertical and horizontal operators are defined by the following relations

vertical:

$$a = \{(k, l) | k = 1, -p \,\square\, l \,\square\, p\}$$
$$A = \{(m, n) | m = 1, -q \,\square\, n \,\square\, q\} \tag{4.14}$$

horizontal:

$$a = \{(k, l) | l = 1, -p \,\square\, k \,\square\, p\}$$
$$A = \{(m, n) | n = 1, -q \,\square\, m \,\square\, q\} \tag{4.15}$$

$$q > p, \tau = (q - p) \tag{4.16}$$

where a is the small neighborhood of length p and A is the large neighborhood of length q. Using the vertical and horizontal binary images the potential control points are obtained by finding the cross points of vertical and horizontal elongated structures. This is implemented by applying a logical AND operation to the vertical elongated structures image Λ and horizontal elongated structures image Γ yielding Υ image which only contain cross points.

$$\Upsilon = \Gamma \odot \Lambda \tag{4.17}$$

Depending on elongates structure thickness the cross points could contain multiple pixels. In cases like these, the centroid of the group of pixels is defined as the potential control point.

Following the method defined in [7], a Gaussian kernel is passed over the image several times to blur the image in an effort to reduce the effects of fine details in structure detection. This leads to detection of only the most prominent elongated structures. Applying this process to raw images produces an intractable amount of potential control points[7]. Figure 4.17 and 4.18 shows a raw and blur image with their respective elongated structure images.
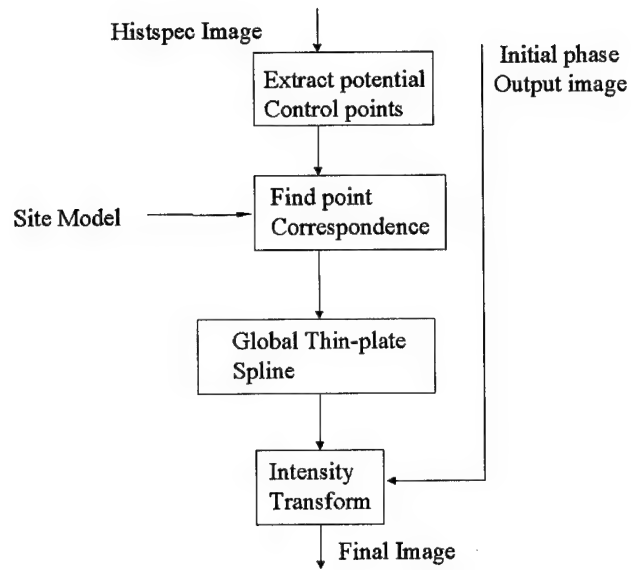
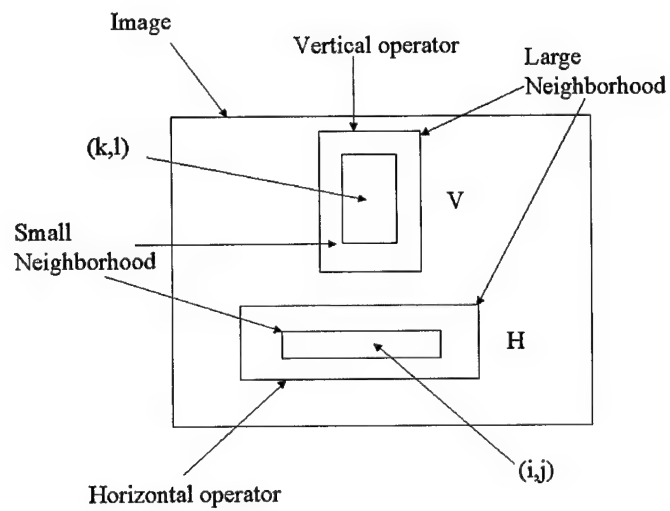Figure 4.15: Process flow for final registration phase
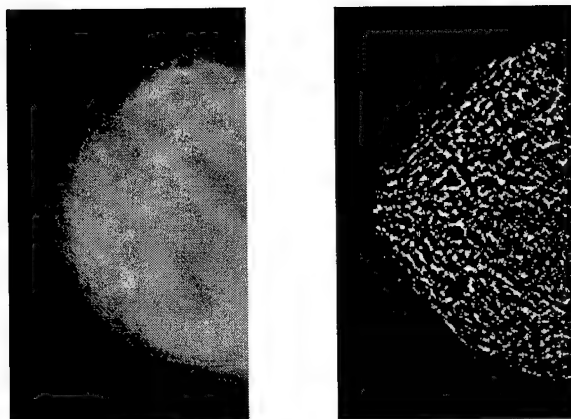

Figure 4.16: Monotony operators for an image

Figure 4.17: Raw mammogram and assocated elongated structures
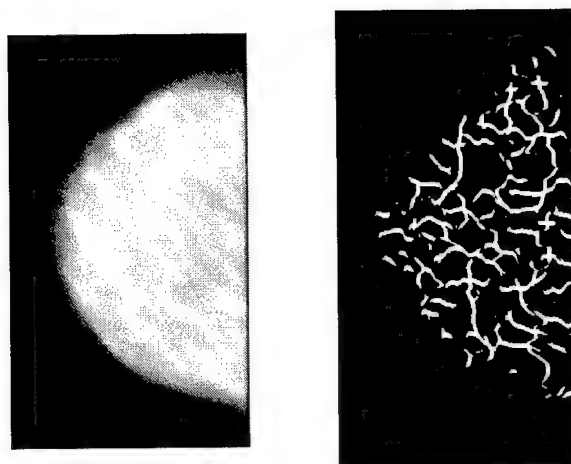

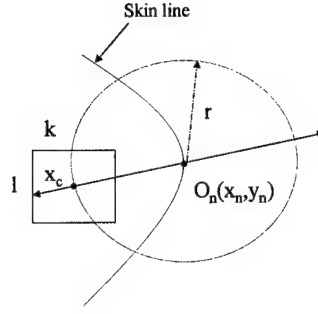Figure 4.18: Three pass filter mammogram with associated elongated structures

Figure 4.19: Matching window location on new mammogram

4.1.10 Point Correspondence

The next step in the fine registration process is matching corresponding control points from the associated pools of control points in each image. Several methods for point correspondence have been investigated and proposed by [7]. These included a signature matching, which is an algorithm that search for longest direction of an elongated structure cross point, and a wavelet based approach that examined localized regions. In addition, [5] used laws texture features to determine correspondence. This research presents two new correspondences methods. The first is based on the signature matching algorithm by [7], but an attempt is made to match the complete structure not only longest direction. The second method transposes the new potential control points $O_q(x_q, y_q)$ onto the old image and matches control points based on point distance from an old potential control point $O_p(x_p, y_p)$. To improve matching rates on both methods, only a subset of the potential control pool from the new image are tested at a single time. This subset is identified as potential control points contained in a $k \times l$ window centered around the point X_c.

The point X_c is the intersection point between a circle centered around the estimated nipple location $O_n(x_n, y_n)$ in the new image and a straight line passing through O_n with a slope of m as shown in below. The slope m of the line is equal to the slope of a similar line in between the potential control point $O_p(x_p, y_p)$ in the site model (old image) and O_o the nipple location in the old image.

$$y = m(x - x_n) + y_n \tag{4.18}$$
$$m = \frac{y_p - y_o}{x_p - x_o}$$

$$(x - x_n)^2 + (y - y_n)^2 = (x_o - x_p)^2 + (y_p - y_o)^2$$

Figure 4.19 shows a pictorial example. Next, each correspondence method will be discussed.

4.1.11 Elongated structure matching

After passing the location criteria ($k \times l$ window), signatures for each potential control point contained, in the local window, are calculated. The signatures are designed to capture the characteristics of the elongated structures surrounding a potential control point. The signatures are calculated by forming the elongated structure image which contains both vertical and horizontal structures. This is realized as a logical OR operations on the vertical and horizontal structure images as shown below.

$$\Omega = \Gamma \oplus \Lambda \tag{4.19}$$

The image Ω now contains cross points and associated vertical and horizontal elongated structures. Figure 4.20 shows some elongated structures derived from a mammogram.
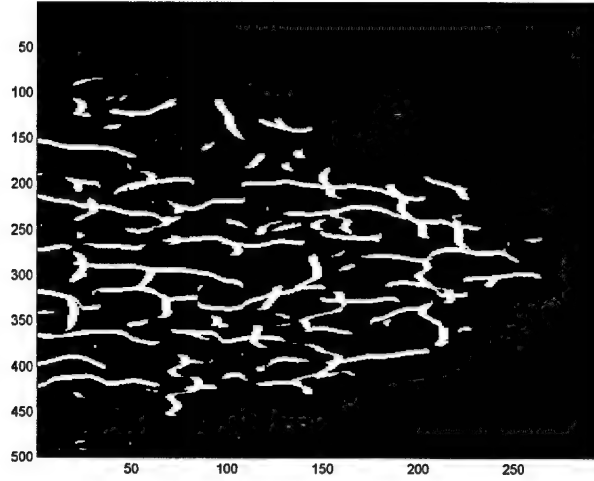
Figure 4.20: Elongated structures detected by monotony operators

The next phase of signature construction is the rotation of a $m \times n$ window N_s steps around the control point. This yields $\Delta\xi°$ for each step. For each step the number of nonzero pixels (NZ) contained within the sum window are counted. The number counted for each step is the signature $y(\Delta\xi°) = NZ$. This process is shown in Figure 4.21. The signatures are then matched by measuring the Pearson correlation coefficient [14] between a pair of potential control point signatures. The resulting coefficient is then applied to a threshold. The Pearson correlation coefficient is formulated by the follow equations

$$\rho = \frac{SS_{xy}}{\sqrt[2]{SS_{xx}SS_{yy}}} \tag{4.20}$$

$$SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{N_s}$$

where y is the N_s point signature of O_p. The Pearson coefficient measures the statistical distance of two distributions. Because non-rigid deformation occurs between images the corresponding control point signature could be a circularly shifted version of each other as seen in Figure 4.22. To consider this problem, the complete signature of the new image control point is circularly shifted by one sample and then Pearson matched. The highest Pearson between all shifts is taken to be the resulting Pearson value for that (O_p, O_q) pair.

The Pearson results for a (O_p, O_q) pair are stored in a modified accumulator matrix. The accumulator matrix is a $N_o \times N_n$ matrix where N_o and N_n are the number of potential control points in the site model (old) and new images respectively. In traditional accumulator formulations [7] ??, the element (O_p, O_q) is incremented each time point O_p matches point O_q, but in this research we put the maximum Pearson correlation coefficient the element corresponding to (O_p, O_q). The final match is performed by taking the maximum value down the columns and zeroing the other column entries for that column. This is followed by taking the maximum value in each row and zeroing the other row entries. The resulting matrix should contain only one nonzero value per row and column. The nonzero elements are the control points.

4.1.12 Simulation experiments

Pearson based control point matches were obtained for the phantom and several real mammograms. The phantom sequence was composed of two versions of the same image. The second image in the sequence was a rigidly transformed copy of the first image. The real sequence contained two images of the same patient acquired at different times. Figure 4.23 shows the potential 'o' and real control points '*' for the phantom sequence where 37 out of the 43 potential control points where matched across the sequence. Compare this to Figure 4.24 where only 5 out of the 36 potential control points where matched. This difference in final control point matching is the result of the variability of extracting elongated structures from mammograms. In Figure 4.23, the structures remain
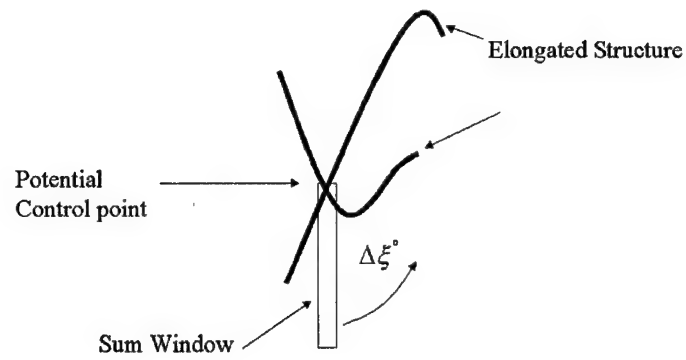
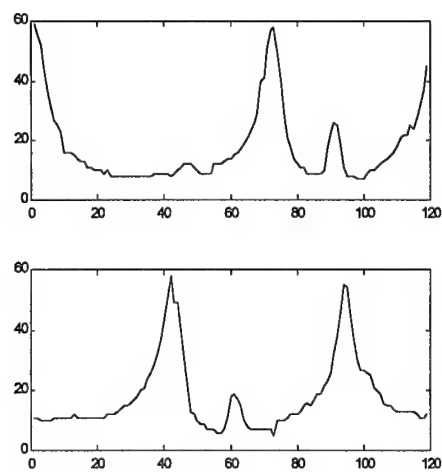Figure 4.21: Formation of potential control point signature.

Figure 4.22: Potenial control point signature with corresponding shifted version
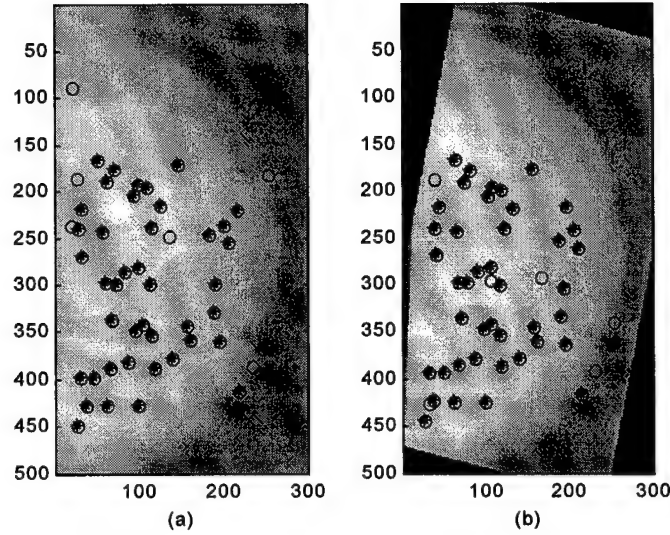
Figure 4.23: Potential and Matched Control points via Pearson matching for the phantom study

stable because the rigid transform causes the signatures to be rotated versions of each other which allows for easy matching. But in Figure 4.24 non-rigid deformation between the image causes the signatures of potential control points to look drastically different if detected at all. In [58], which uses much the same approach but only considers a 40×40 window using the longest arm of the structure as the matching metric, only obtains 6 control points for a real sequence. In this research, a smaller 10×10 window is used along with the Pearson matching criteria to obtain comparable results. This reduction in window size is attributed to use of the complete signature information in matching not just the most dominate structure arm. To increase matches, the local match window currently at 10×10 should be increased. It should be noted that this operation also increases false match probability and processing time.

4.1.13 Nearest Neighbor match

In this method, initial registration is assumed to have corrected most of the global distortion and mis-adjustment between the two images. The control point correspondence is then obtained by overlaying the potential control points from the new image with the potential control points of the old image and calculating the Euclidean distance from each old potential control point to each new potential control point.

$$d_j = \sqrt[2]{(x_i - x_j)^2 + (y_i - y_j)^2}$$

with i and j equal to the index of potential control points bounded by $i = 1......Nold$ and $j = 1......Nnew$. The new potential control point with the smallest d value is selected as a match for the old point of interest. Figure 4.25 shows a typical case of a localized window. In the event, a new potential control point is matched to several old points the match with the smallest d is selected as the final match.

4.1.14 Simulation Experiments

Figure 4.26 shows the same sequence shown in Figure 4.24 where nearest neighbor matching is used. This matching methodology more than doubles the number of matched control points over matching with Pearson matching method. It also produces control points that are distributed evenly around the image. This method exceeds the method presented by [7] at smaller matching window sizes. A key note is the dependence of this method on initial registration. Without initial registration, distance is not a good enough metric along. Again more matches can be obtained by increasing window size at a cost to processing time and false match rate.

4.1.15 Spatial-coordinate

The main goal in registration is to obtain a transform T_A such that one of the images could be transformed into correspondence with the other. In general, an image mapping transform is represented by
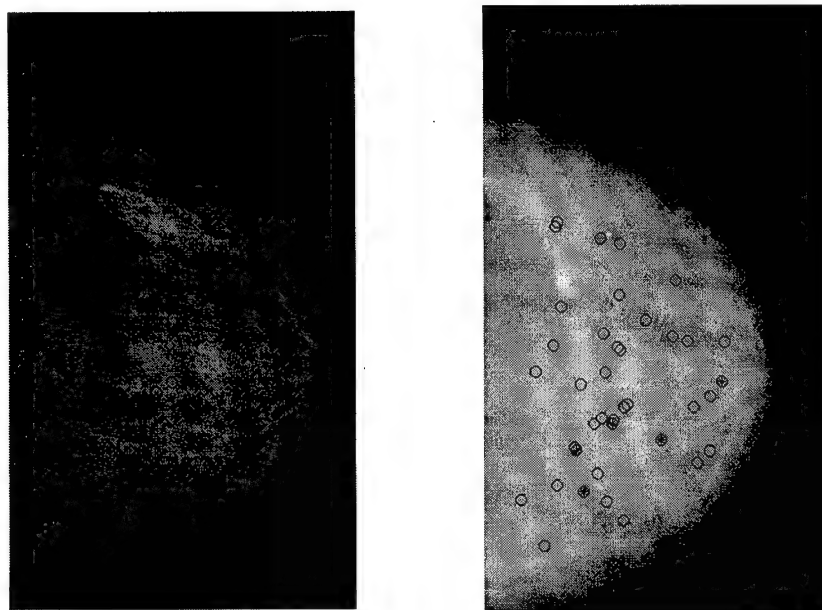
Figure 4.24: Potenial control points shown by o and matched control points shown by $*$ via Pearson matching
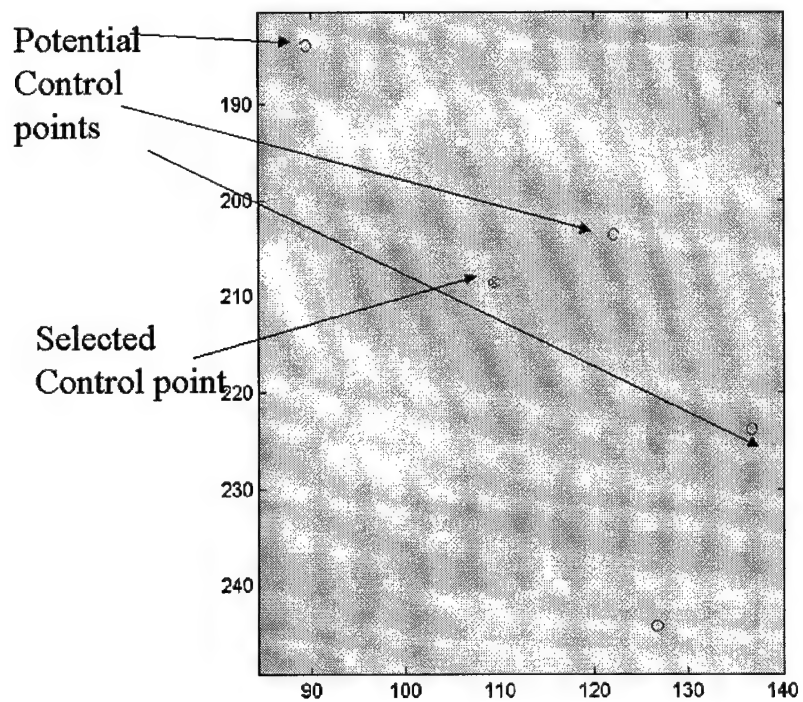

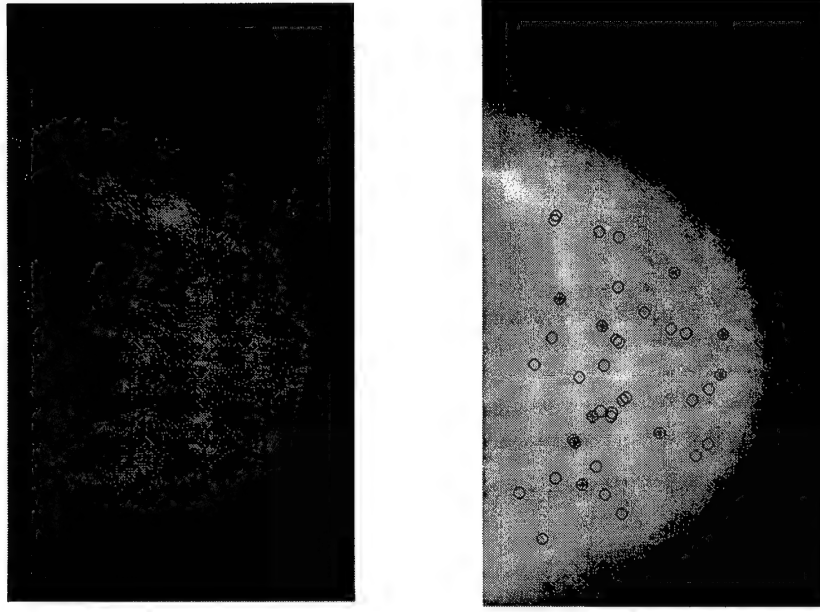Figure 4.25: Local correspondence window for a potential control point

Figure 4.26: Potenial control points shown by o and matched control points shown by $*$ via Nearest nieborhood method.

$$T_A(x,y) = (f_x(x,y), f_y(x,y)) \tag{4.21}$$

where $f_x(x,y)$ is the mapping function for x coordinate of (x,y) and $f_y(x,y)$ is the mapping function for the y component of (x,y). Since breast tissue is inherently nonrigid, complex changes can occur between the image in the sequence. To account for these changes, the function $f()$ needs to be non-linear. [5],[68] selected TPS as the mapping transform so we apply it in our case. The mapping function for TPS is shown below

$$f(x,y) = w_0 + w_1 x + w_2 y + \sum_{i=1}^{n} W_i g(r_i) \tag{4.22}$$

$$g(r_i) = r_i^2 \log r_i^2$$

given that $r_i = (x_i - x)^2 + (y_i - y)^2$. This transform is made up of a global (affine) portion and (elastic) portion. These two portions are distinct but can be evaluated simultaneously.

In order to use $f(x,y)$ to transform the image, the coefficients w_0, w_1, w_2, W_i must be estimated. This is done by using the control points determined form the previous section, to formulate a least square approach to coefficients estimation. The least squares formulation starts with coordinate mapping relation

$$(u,v) = (f_x(x,y), f_y(x,y)) \tag{4.23}$$

where (u,v) is a point in the new image (control point) that is associated with the point (x,y) in the old image (control point). Given (u,v) and (x,y) are control points, zero error should occur when transforming (x,y) through the mapping function.

$$(u,v) - (f_x(x,y), f_y(x,y)) = 0$$

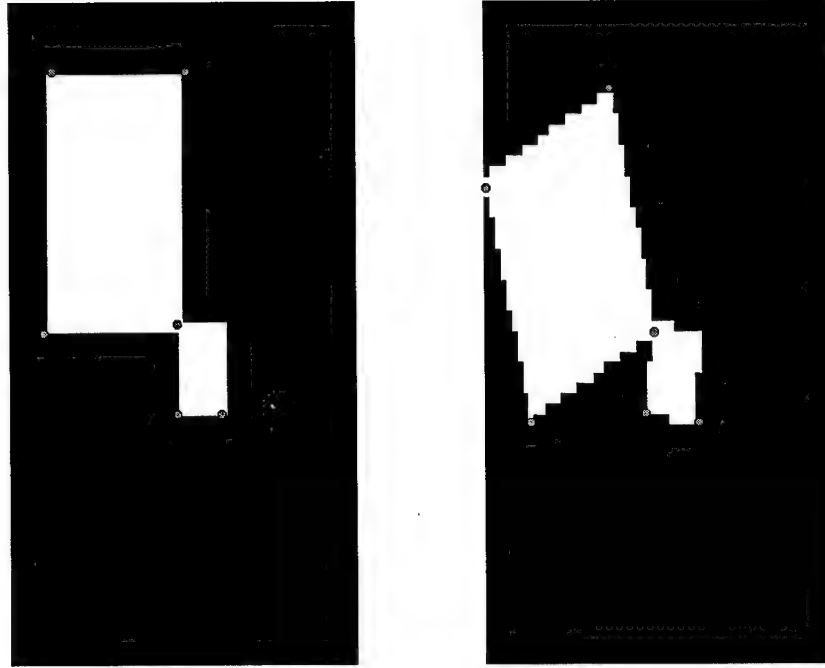Rearranging terms and expanding to handle n control points a general error equation is formed given below.

Figure 4.27: Raw phantom sequence

$$E = \sum_{i=1}^{n} \left[(u_i - f_x(x,y))^2 + (v_i - f_y(x,y))^2 \right] \tag{4.24}$$

The above equations leads to the normal equations. The relation for the x mapping functions is shown below.

$$\sum_{i=0}^{m}\sum_{j=0}^{i} a_{ij} \left[\sum x_k^j y_k^{i-j} x_k^\beta y_k^{\alpha-\beta} \right] = \sum_{k=1}^{n} u_k x_k^\beta y_k^{\alpha-\beta} \tag{4.25}$$

where $\alpha = 0....m$ and $\beta = 0....\alpha$. The coefficients for the y mapping functions are found in a similar fashion. With the mapping functions f_x and f_y each pixel is then transformed to produce the warped image. In general, the new pixel location will not fall on a exact grid point some interpolation is used to obtain the pixel value. In this research, nearest neighborhood interpolation is used to determine the new pixel value.

4.1.16 Simulation experiments

This process is examined through the following example of a phantom that is made up of two squares where each square is transformed by a different amount. The image pair is shown in Figure 4.27. Table 4.4 shows the mse between the reference and the stages of the warped image. From the table one can see the mse decrease through out the process. Use of PAR along reduces the mse by 77%. With the addition of TPS the mse is reduces by another 10%. A small decrease in mse after PAR is attributed to the use of only 6 control points. If more control points had been selected the performance gain of TPS in this process should improve.

4.2 Summary

This registration approach is composed of two main steps an initial step and fine step that are supported by the site model. The site model supports the registration process by storing user (manual) and automatically extracted
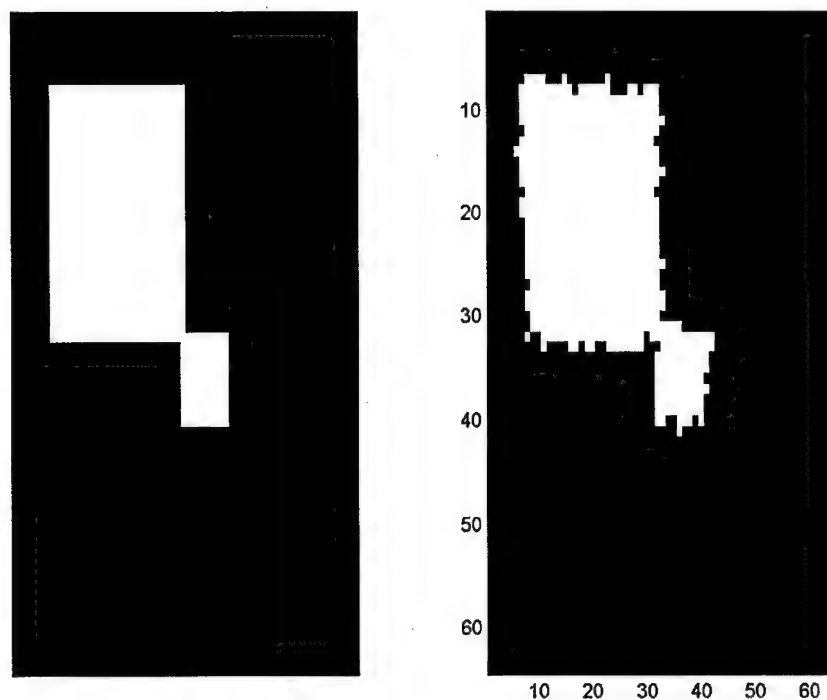
Figure 4.28: Registered phantom sequence

Number in error	Method
395	none
90	PAR
60	PAR-TPS

Table 4.4: Amount of pixels in error for registrations methods

input for use in registration. The model also provides a common frame for incoming images to register to. Finally, the site model stores the complete image sequence history in a common place. The initial registration step is aimed at addressing gross misalignment between the images. This step is rigid model deformation based and requires little environment knowledge (i.e. control point locations). While the fine registration step requires the identification of corresponding control points. The fine step is aimed at correcting non-rigid deformation between images in the sequence. Together mammograms can be robustly registered in support of change analysis. With the mapping functions derived above each pixel is transformed to produce the new image.

Chapter 5

Site Model Supported Change Detection

5.1 Introduction

Change detection is the process of identifying significant differences as measured by a metric between two or more objects. In this research, the objects of interest are images or sub-images (i.e. localized windows) in a sequence. In an image sequence with objects, three types of change can be defined. In the first type of change, defined as type I, only intensities of the pixels change. In the second type, defined as type II, the intensities remain constant, but the location or shape of the object changes. In the third type of change, defined as type III, intensities, shape, and location change. These types of change can be measured either pixel by pixel or image by image. A simple formulation of a pixel change metric is shown below.

$$D = \Im(R_f, R_r) \tag{5.1}$$
$$image(i,j) = \begin{cases} 1, & D(i,j) \propto \gamma \\ 0, & D(i,j) \propto \gamma \end{cases}$$

where D is a change map containing the metric measurements at each pixel. $\Im()$ is the pixel function criteria applied for processing. For example, in difference analysis the function $\Im$ would equal abs. γ is the metric threshold, R_f is the transformed image, and R_r is the reference model image. Image change is measured in much the same way as pixel, but the image is evaluated as a whole.

$$D = \Im_O(R_f, R_r)$$
$$mage = \begin{cases} 1, & D \succ \gamma \\ 0, & D \prec \gamma \end{cases} \tag{5.2}$$

where $\Im_O$ is the image change function, D is a scalar change value, R_f is the float image, and R_r is the reference image. An example of an image change function could be the mutual information between to image blocks as shown below where

$$\Im O = \sum p_{xy} \log p_{xy}$$

p_{xy} is the joint distribution of an image x with marginal density p_x and an image y with a marginal density of p_y.

Change detection in images has found application in various fields including video sequence processing; satellite imaging; and medical imaging. In video sequence processing, numerous change detection metrics have been developed [10]. The main goal in this application is to find abrupt scene changes to aid in sequence compression. The compression is achieved by sending only a reference image (i.e. first image in sequence) then only scene change information (global) in subsequent transmissions. The video change metrics assumes high SNR and the occurrence of abrupt change. The main motivation is to detect the region of the image that contains the change. No effort has been put into describing the change. The most research on change detection has been conducted in the satellite imaging area (remote sensing). In this area, work has been done on building change detection, agriculture crop analysis, and weather tracking [79], [82]. Some specific change metrics have been developed for synthetic aperture (SA) images [83], but they take advantage of the multi-spectral data that is inherent to SA imaging. For this reason, they are not as useful for other applications (i.e. non-SAR applications). Again, as in video change, no effort has been put into describing the type of change.

In the medical environment, the existence of change and the classification of change are very important. This change leads to valuable diagnosis information. Since the change metrics for video requires high SNR and the metrics for SRA are SAR signal dependent, a new metric is needed. The newly developed change process should also have
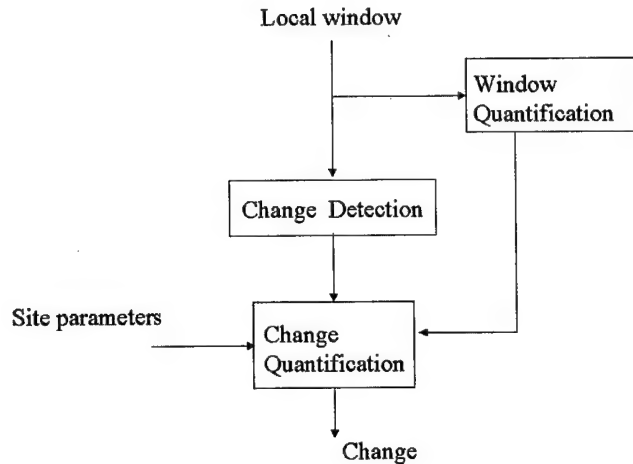
Figure 5.1: Change detection process flow

the capability to quantify change. To accomplish this task, a two step process is developed. The steps are change detection and quantification. The detection phase is performed by measuring the joint relative entropy between the two objects with entropy values higher than a user specified threshold marked as change. Quantification consists of comparing the objects area and center of gravity. Often, in medical applications such as lesions monitoring, change overtime is of great interest as it can show response to drugs or disease progression. The site model, which is a dynamic mathematical and geometrical description of a scene under analysis, has been shown to be a useful tool in the analysis of changing images in a sequence [79]. When applied to the medical change problem, the site model could store the behavior of an object in the scene as well as over all image behavior. Use of the site model also allows the integration of user supplied input (domain knowledge) with automatically extracted parameters towards the goal of change detection. This idea of user input models the real process that a radiologist uses to analysis the image. Specifically, the site model supports change detection in two main ways. First, the site model provides a unified location to store change that has been occurring overtime. This feature is useful in monitoring application. Second, the site model can be used to determine which part of the image should be considered in processing. For example, the changed portion of a image might not be included in transform calculation. This generates a more robust transform. The rest of the chapter considers the development of the change algorithm. The complete block diagram is shown in Figure 5.1.

5.2   Change Analysis Theory

The site model supported change detection algorithm contains two main phases. Phase I is change detection and phase II is change quantification. Change detection is the process of determining whether two objects (images) differ. In practice, nothing is ever exactly the same, so the change detection results are measured in comparison to a threshold. For this research we selected the use of relative entropy as our change metric. Relative entropy is a measure of the inefficiency of assuming that one distribution exactly matches the other. (i.e. distance apart) Relative entropy is given by the equation.

$$D(p//q) = \sum p(x) \log \frac{p(x)}{q(x)}$$

where $p(x)$ and $q(x)$ are the distribution of image P and Q respectively. Relative entropy is also known as Kullback Leibler distance. To utilize this relationship, the distribution of each image is required. These distributions are modeled by the gray level histogram of the image. The resulting $D(p//q)$ valve is then compared to a threshold for change determination. The threshold is selected manually and is highly dependent on image dynamic range. Since spatial information is thrown away during the calculation of the histogram, the use of the marginal densities makes the metric insensitive to type II changes. To address this problem we, consider the use of the joint densities because these densities maintain spatial information. This leads to the formulation of a new detection metric relative

Change type	Orignal object			Change object		
	x, y	size	Intensity	x, y	size	Intensity
III	$205, 205$	10×10	100	$205, 205$	10×10	100
III	$100, 100$	10×10	100	$100, 100$	10×10	100
II	$50, 435$	10×10	100	$58, 426$	10×10	100
I	$250, 250$	10×10	100	$250, 250$	10×10	115
none	$135, 333$	-	-	$135, 333$	-	-

Table 5.1: Configurations of change blocks in phantom.

	q_c	q_d	$\hat{q}_c$	$\hat{q}_d$	GRE	AHST	Chi
1	+300	0	302	0	10.8	.05	.2565
2	0	12.04	9.89	15.86	4.99	.2	.3
3	0	0	0	0	2.87	0	0
4	100	0	-100	N/A	2.56	.0013	.00319
5	N/A	N/A	-	-	0	0	0

Table 5.2: Change Quantification results

entropy.

$$D(p_{xy}//p_{xx}) = \sum p_{xy} \log \frac{p_{xy}}{p_{xx}}$$

This metric measures the inefficiencies of assuming that p_{xx} is the distribution for p_{xy}.

The next phase of processing is change quantification. In this process, the characteristics of the change are determined (i.e. amount, shape, change). This is performed in a multistep process. First, segment the image into two classes. Second, compare the segmentation image with the reference segmented image. Third, form objects from each image and calculate object shape area and center of gravity. Finally, calculate the object overlap and size of difference. The results are then stored in the site model for the next stage of processing.

5.3 Simulation Experiments

To simulate this portion of the system, a phantom mammogram sequence containing four manually changed regions was processed. The three types of change were simulated by modifying a $N \times N$ block of manually changed pixels. Table 5.1 shows the four different configurations. To make the blocks more natural, Gaussian filters are applied to smooth out the edges. To isolate the change detection performance, the phantom sequence was assumed to be perfectly registered. This is accomplished by using the same mammogram in both images of the sequence. We further assume that the radiologist has identified the regions of interest, a 30×30 block of pixels, a pori. Generally, in most change detection metrics a, function is evaluated yielding a value which is then compared to a threshold. For this simulation it is assumed the detection threshold is predetermined at 0.5. The performance of joint global relative entropy (GRE) will be compared to two video sequence metrics, an absolute histogram (AHST) and chi square metric (CHI). The quantification portion will be tested by quantitative comparison of the phantom blocks.

Table 5.2 contains the results from processing the phantom where q_c and q_d are the true Δ area and location respectively; and $\hat{q}_c$ and $\hat{q}_d$ are the estimated Δ area and location. For the detection phase of processing we see that GRE metric obtains favorable detection results on all three types of change. The GRE values are $>>$ than the threshold. This indicates that possibly the threshold can be increased which would improve robustness by decreasing the possibility of noise being flagged as change. On the other hand, AHST and CHI fail to detect change at all. This is attributed to the dependence of these metrics on the marginal densities which do not store spatial information. The values produced by these two metrics are $<<$ than the threshold. One would tend to think that performance for these metrics could be improved by decreasing the threshold, but this would only serve to flag noise differences as change. The superiority of the GRE metric can also be seen by examining the ranges of values. The GRE ranges from 0..10.8, while AHST and CHI teams range from 0..0.3 and 0..0.2 respectively. These ranges can also be called dynamic range (value ranges). In communication systems dynamic range is a indicator to the systems sensitive. This same ideal applies to the detection metrics. The GRE metric has a larger spread than AHST and CHI which allows it to capture more and smaller amounts of change.

In the quantification phase, the algorithm accurately quantifies type III change. In this example, the true area difference was 300 pixels^2. The estimated area difference was 302 pixels^2. In this case, the translation was estimated

with exactly 0 pixels. In the type II change example the areas remained the same, but a translation of 12 pixels was recorded. The algorithm estimated an area change of 9 pixels^2 and a translation of 15 pixels. The error in the area could be attributed to the inability of the object selection process to extract the object. Generally, this occurs when the block is the same intensity level as the background. In type I change, the algorithm estimates 0 area change and 0 translation. To fully test the algorithm, an example was selected were no change occurred at all. These results are shown on the bottom row of Table 5.2. Here we see that GRE, AHST, and CHI did not flag this region as changed, but it is difficult to tell if AHST and CHI really found no change or are producing values in there dynamic range.

Chapter 6

Experimental Results and Discussion

6.1 Introduction

The main objective of this research is to detect biological change in a temporal sequence of mammograms. Different types of change can occur between mammograms acquired overtime. The first type of change is natural change which includes weight change and tissue composition change. The next type of change is image acquisition change. This includes the changes caused by breast positioning, breast compression, and differences in imaging equipment. Finally, change that possibly indicates cancer or the onset of cancer. This type is usually visualized as a microcalcification or mass [3]. The first two types of change generally affect the complete image and are classified as global change. On the other hand, the third type of change is usually localized to a region and is classified as local change. Due to the enormous number of combinations relating to the first two types of change, we focus attention on local change. In addition, we also only consider change calculated from a radiologist selected localized window. Local change has been shown to be an indicator of the onset of cancer [4]. Currently, radiologists perform change analysis manually following a specific procedure [3]. Automation of this task could help to reduce the fatigue felt by the radiologists which may lead to an increase in analysis accuracies. This chapter presents and discusses the results generated by applying the developed change detection algorithm to real mammogram sequences. See Figure 6.1 for a system overview and flow diagram. Next, the results of several example mammogram sequences will be discussed.

6.2 Experiment Results and Discussion

The first example is a sequence composed of two right CC views of the same patient acquired on 1/21/93 and 2/3/99 as shown in Figure 6.2 a and b. The image acquired on 2/3/99 contains a suspicious region located at $(77, 317)$. Figure 6.2a is taken as the reference image and used to construct the site model. The users input to the site model is the region of interest, which is a 30×30 square centered around the point $(77, 317)$. The radiologist selects the window size manually as seen in Figure 6.3. After construction of the site model, processing new images can commence. The first step is the extraction of parameters used in initial registration. This includes objects and their descriptions. Next, multi-object PAR is performed using 2 of the objects as seen in Figure 6.4. The resulting initial registration pair is shown in Figure 6.5. Comparing Figure 6.5 and Figure 6.2 we see that most of the scale difference between the images has been corrected. Finer alignment could be obtained if control points were known. Using the initially registered image, final registration parameters are extracted. These parameters include potential control points and their associated signatures. Next, the recently extracted potential control points are matched with the potential control points from the site model to obtain the final control points. This matching is performed by two methods in this research. Figure 6.6 shows control points obtained by matching signatures using the Pearson correlation coefficient while Figure 6.7 shows control points obtained by matching Nearest Neighbor. In this example, Pearson matching yields 13 control point pairs out of a pool of 66 potential control points or a match rate of 0.197. This rate is low because the deformation between the site and incoming image produced different potential control point pools in each image. Thus, signature matching yields few matches when signature correlation is low. The final control points in this example are clustered into 2 loose groups located on the top and bottom of the breast. This appears to be caused by the existence of dense tissue near the center of the breast. In dense tissue, the monotony operators (used to find elongated structures) appear to have problems when the tissue intensities are nearly constant. Nearest Neighbor matching, on the other hand, yielded 27 control points evenly distributed across the image. This yields a match rate of 0.409. This number is still low, but more acceptable. Both matching rates could be improved by the increase in the localized search window size, but the probability of mis-match would also increase. Mis-match control points cause gross distortion in the transformed image. Since our method of control
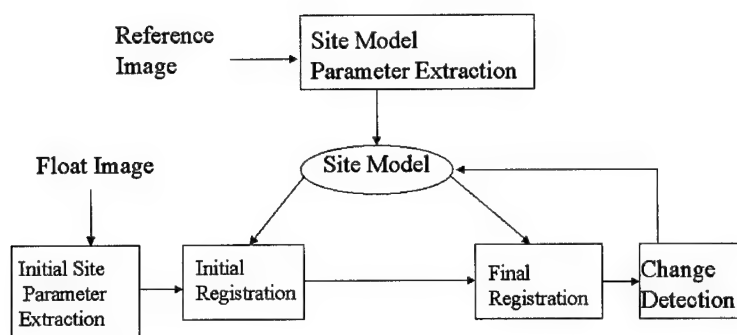
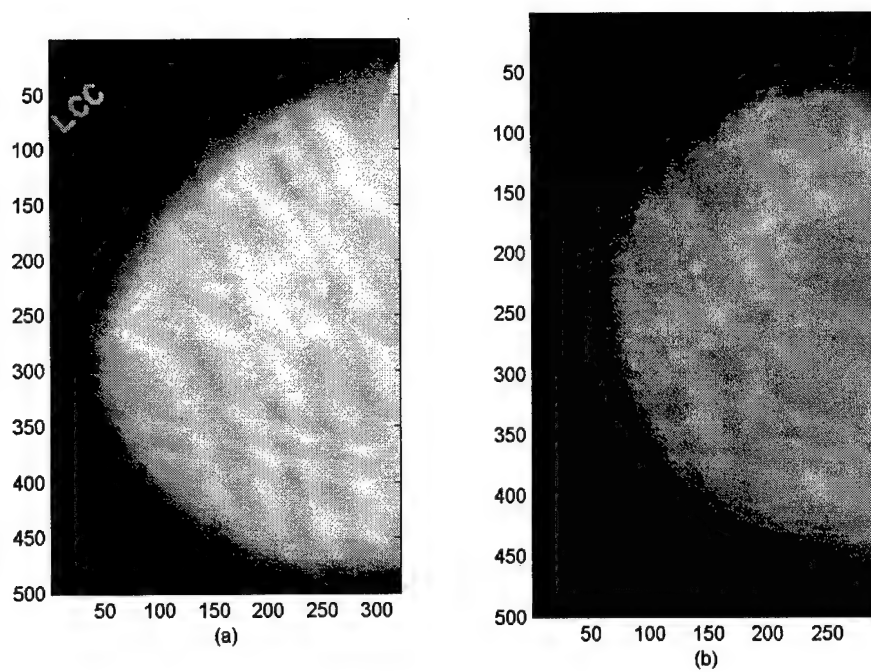Figure 6.1: Change detection process flow

Figure 6.2: Raw mammogram sequence. (a) 1/21/93. (b) 2/3/99.
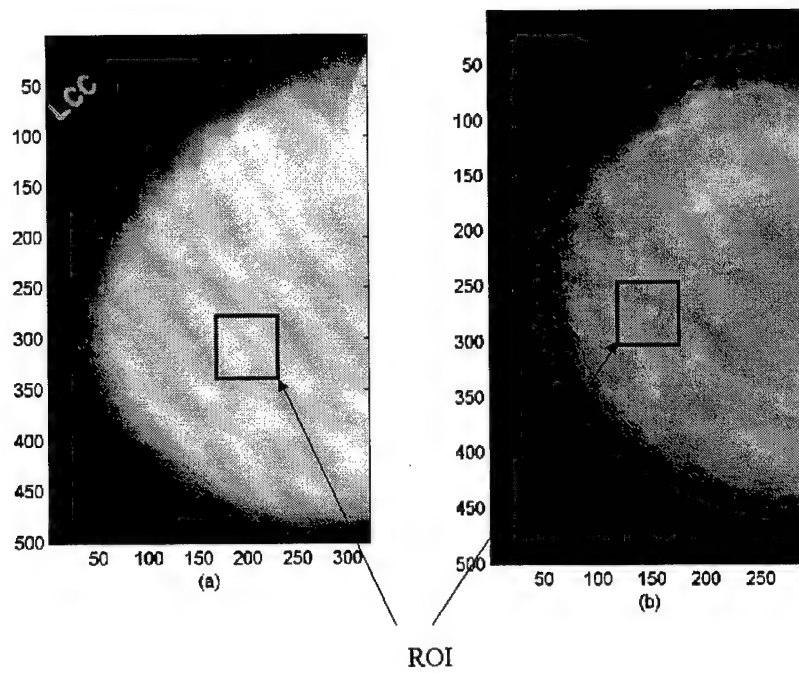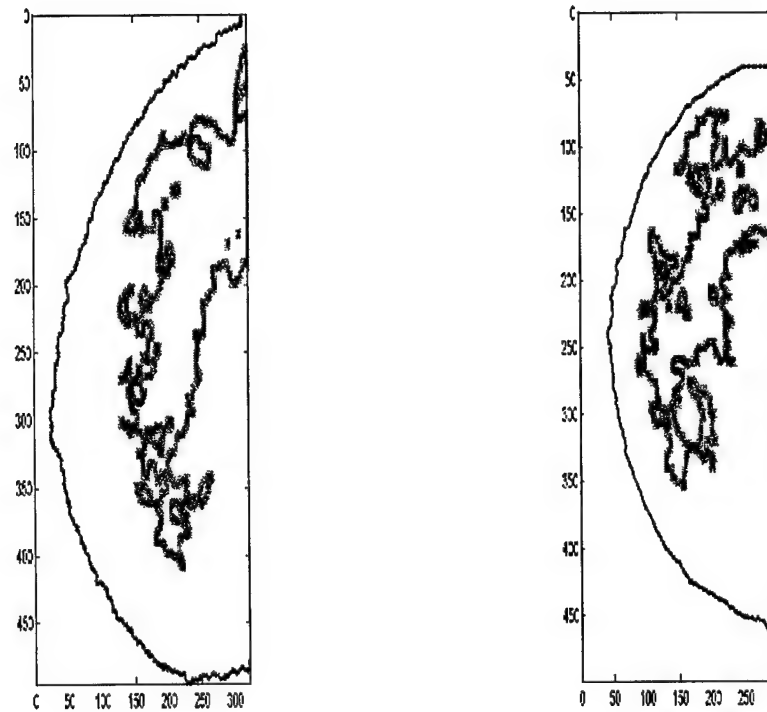
Figure 6.3: Marked region of interest.

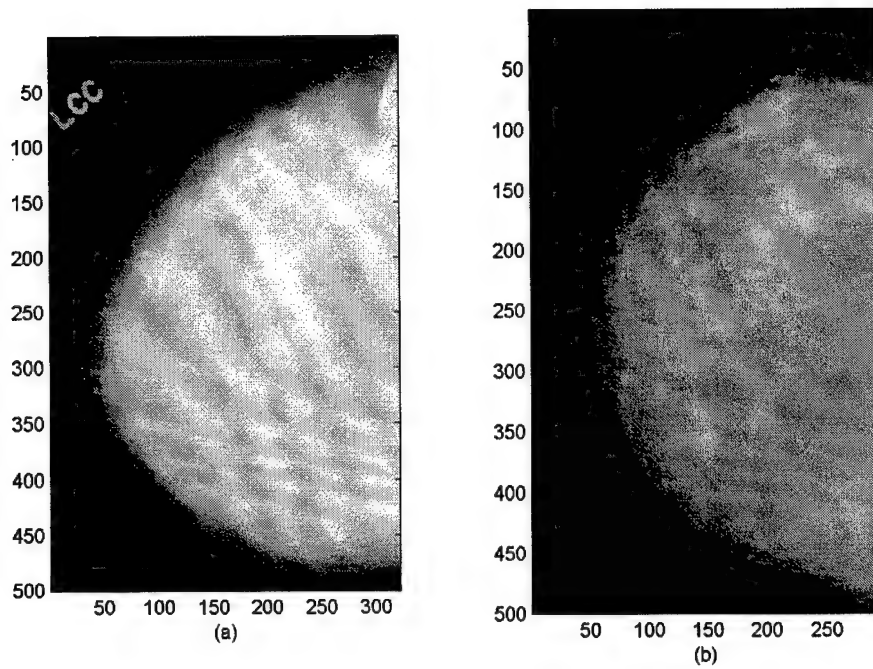Figure 6.4: Objects used in Multi-object transform.
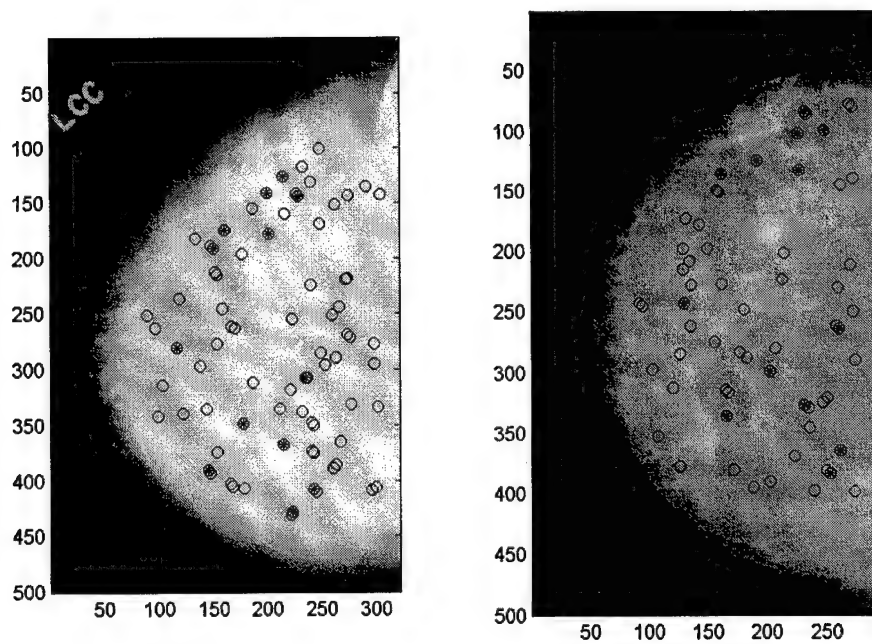
Figure 6.5: Multi-object PAR image pair.

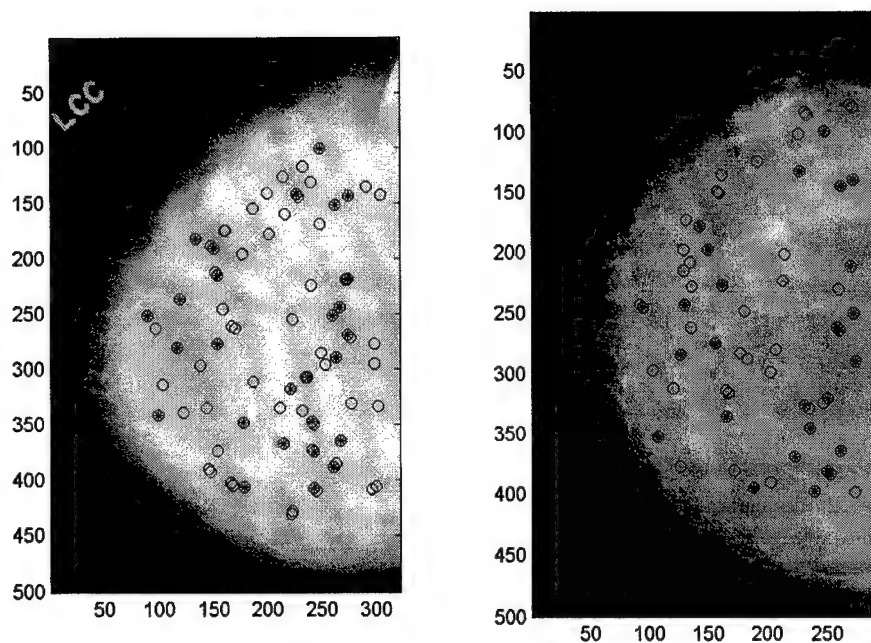Figure 6.6: Potenial 'o' and final '*' control points using Pearson correlation.

Figure 6.7: Potential /o/ and final / * / control points using Nearest Neighbor method.

point extraction is based on [58] we suffer the same image dependence problems as [58]. Window sizes, thresholds, and monotony operator dimensions are among the key parameters that need to be adjusted on a per image basis. For our research we use a window size of 10×10, threshold of 6, monotony operator dimension of 1 and 5. These values were experimentally determined using visual inspection of initial output. Next, the final transform is derived and applied to the image pixel by pixel resulting in the pair shown in Figure 6.8.

To perform change detection, the corresponding region of interest from the incoming image is compared to the site model. The histograms of the two regions are compared in Figure 6.9. From this figure, the difference is visually apparent as the two regions have different distributions. Three change metrics were applied yielding the following results: global relative entropy (GRE) 23.63; absolute histogram difference (AHST) 0.885; and chi square (CHI) 1.0. The last two metrics are video sequence metrics and serve as comparisons of existing change methods. Given the threshold of 1.5 which was determined experimental, both AHST and CHI miss the change which means they appear to be insensitive to slight scene changes, but GRE detects the change. In fact, this change resulted in a GRE value $\gg 1.5$. It would appear that the threshold could be increased, but this would increase the probability of miss.

Unlike the phantom studies performed in the other chapters, no ground truth exists for quantification of the changed region. For this reason, visual inspection is used to examine the results. The quantification process determined an area difference of 353 pixels which was verified by an radiologist during a manual inspection. The detected area is larger then the area estimated by the radiologist because the object extraction process cannot remove all of the background pixels. 54 out of the 354 pixels are background pixels.

In the next example, the radiologist identified a suspected area (region of interest) on the final mammogram (i.e. first image). The raw sequence is given in Figure 6.10 and is composed of a right CC view of a patient acquired on 3/5/96 and 2/24/99. The /X/ marks on the image are the location of the change region. On the site image the /X/ is the associated point. For this example, two objects were selected for use with the multi-object PAR. Figure 6.11 is the resulting transformed image where /X/ marks the change location. ___ control points were matched out of ___ potential control points to form the TPS transform. The final warped image is shown in Figure 6.12. From examination of the image it appears distortion occurred, but the location of the /X/ on both images appear to visually cover the same portion of tissue. In comparing, Figure 6.10, 6.11, and 6.12 we indeed notice this fact. The image's distorted look is caused by too few control points on the skin line (or region). Thus, the affect of the
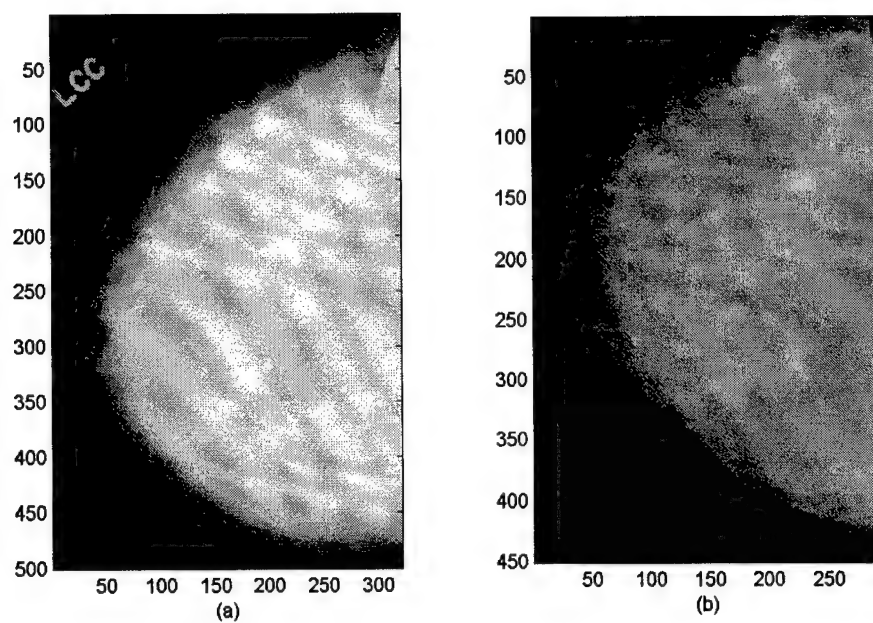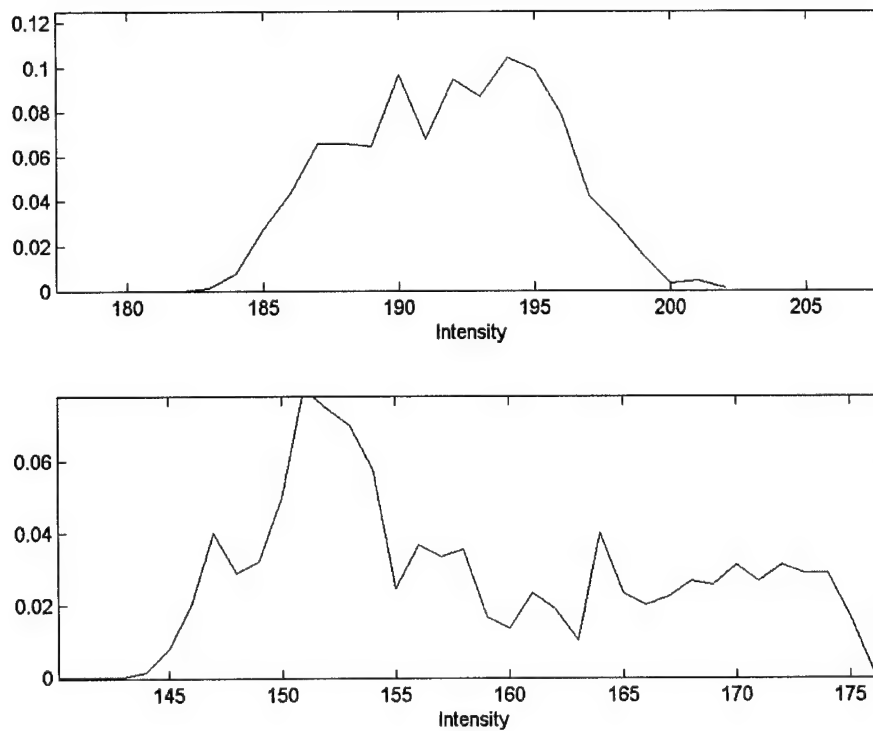
Figure 6.8: Final warped image pair.

Figure 6.9: Histogram of corresponding regions of interest
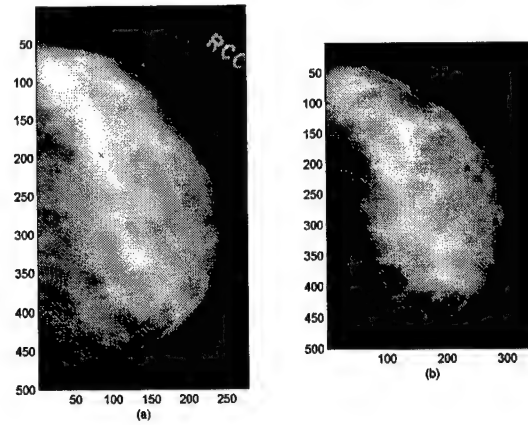
Figure 6.10: (a) Reference image 3/5/96 , (b) float image 2/24/99.
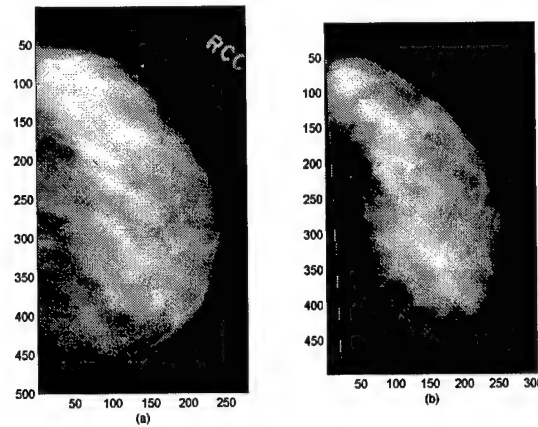

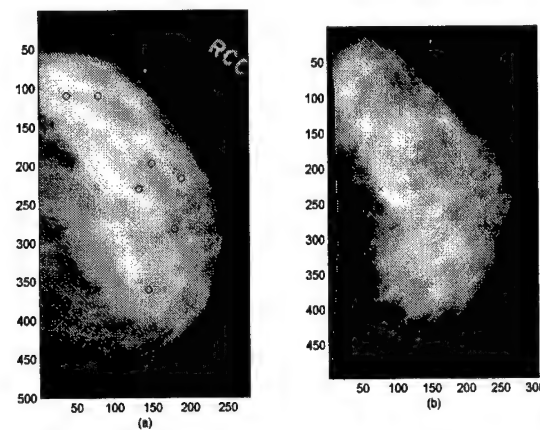Figure 6.11: (a) Reference image, (b) Multi-object PAR image.

Figure 6.12: Final image pair after registratoin. (a) Reference image. (b) Warped image.
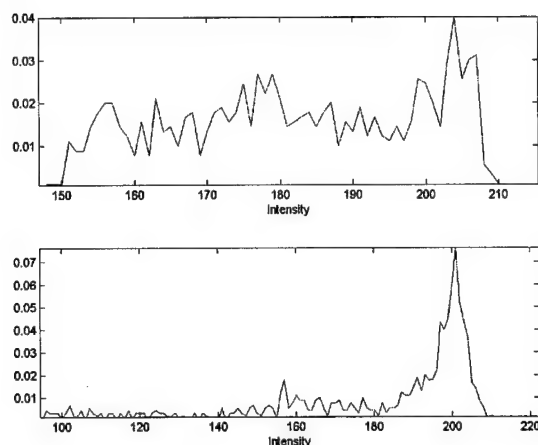
Figure 6.13: Histogram comparison between the two local change windows

control points on the skin-line pixels is greatly reduced causing a massive warping effect. The algorithm was then used to see if the area was present on the site mammogram. The intensity histograms of two regions are shown in Figure 6.13. From here change can be visually determined. To detect the change, GRE, AHST, and CHI metrics were calculated yielding the following values 22.9, 0.512, and 0.4611 respectively. Again, the GRE metric is $\gg$ then the threshold while AHST and CHI fall below the metric. The quantification results estimate a 530 pixels^2 change. The true change is closer to 9 pixels^2. The massive error results from the inability to extract the object from the background of similar pixel intensities resulting in a large selected region.

6.2.1 Summary

Change detection not only highlights existence of possible changed regions, but when combined with the site model provides a patient history by showing site progression. One of the key components of change detection is image registration. In this chapter, we applied our multi-step registration algorithm to mammogram sequences. Acceptable registration and change detection were obtained. Improvement in control object selection and control point extraction would go along way to improving the overall results. The key to registration is landmarks between the images. In this research, we use objects and points as landmarks. Current methods of object and point selection are image dependent and adhoc. Incorrect assignment of control points/objects could cause erroneous transformation. This change detection is not exact, but would be sufficient to flag a radiologist to review the area. The main results of this study consisted of the automatic alignment of mammograms, detection of change in a local window, and implementation of a mechanism to store and build up patient information via the site model.

Chapter 8

Appendix A: Information Criterion

Determining the number of components in a mixture signal is useful in numerous applications from speech processing to object recognition. These type of problems are termed model selection or cluster validation in the literature [23]. The main goal in these type of problems is to estimate, given the data, the number of components K, are present in the mixture signal. This is accomplished by evaluating a function (Information Criterion IC) for reasonable valves of K. $\hat{K}$ is taken as the K value that yields the minimum function result. The first and most widely used IC is Akaike Information Criterion (AIC).

8.1 Theory

The AIC formulation can be derived using the following model [23]. Suppose our data is represented by N random vectors given by $Y = \{y_1, ...y_N\}$. Further assume that the distribution of y is composed of K components where the distribution of the kth component is $f_k(Y/\theta_{ml}^k)$ where θ_{ml} are the ML estimate of the features. So the goal of the IC is to find the K that maximize the function. Since we assume our distribution is a Gaussian, finding its maximum is equivalent to minimizing the log of the distribution function. The results are the AIC equations given below.

$$AIC(K) = -2\log(f(x/\phi_{ml})) + 2 * K_a \qquad (8.1)$$

$$K' = \arg\min AIC(K); 2 \preceq K \preceq K_0 \qquad (8.2)$$

where $f(x/\phi_{ml})$ is the conditional likelihood function distribution given the maximum likelihood feature vector ϕ_{ml}. K_a is the number of free parameters to estimate and was added to make the AIC estimate an unbiased estimate of the mean distance between $f(x/\theta)$ and $f(x/\theta')$ where θ' is the estimated parameter vector.

8.2 Simulation Experiments

To illustrate this algorithm two examples were processed a four class phantom shown in Figure 8.1 and a real mammogram. For each example, the k ranged from 2..10. Figure 8.2 shows the plot the AIC curve for the phantom and Figure 8.3 shows the plot for the mammogram. From these plots we see that $\hat{K}$ is 4 and $\hat{K}$ is 8. The results correspond to results achieved in [27].
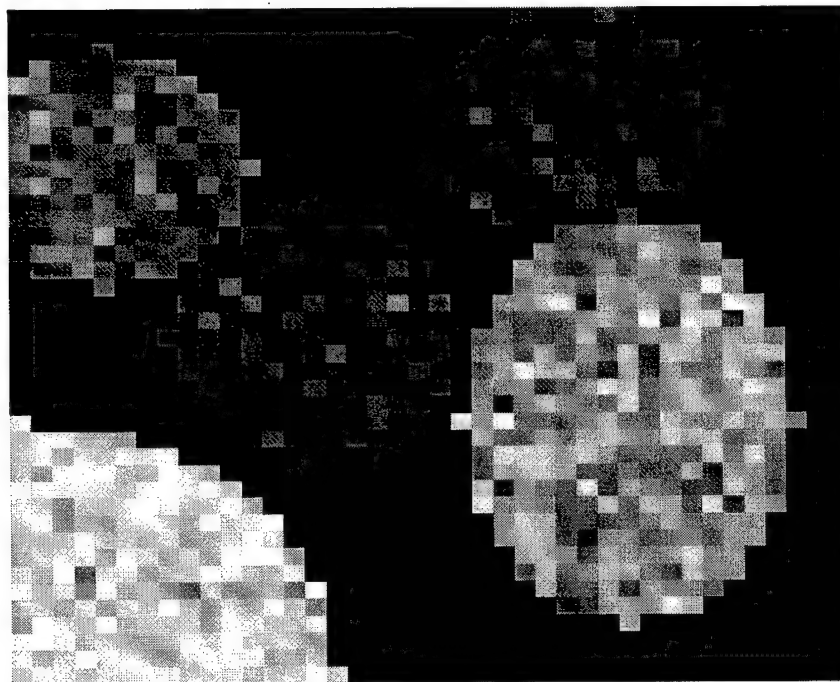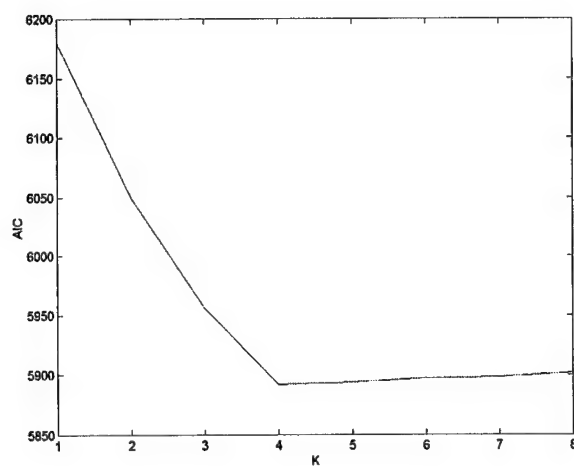
Figure 8.1: Four class phantom
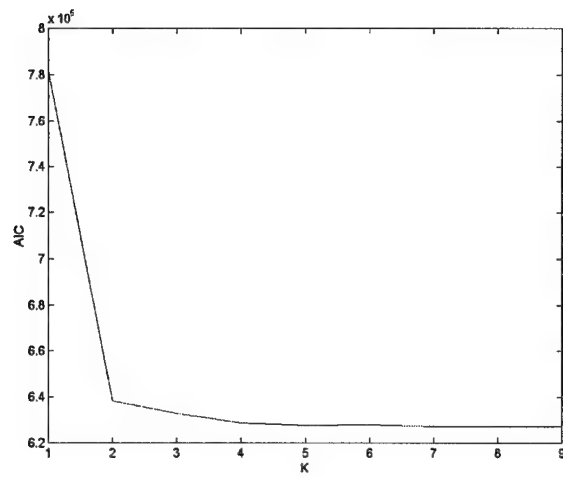

Figure 8.2: AIC plot of four class phantom

Figure 8.3: AIC plot of mammogram image.

Bibliography

[1] D. B. Kopans, *Breast Imaging*, J.B. Lippincott Co.,Philadelphia, Pa., 1989.

[2] D. Dance, "Physical Principles of Breast Imaging," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp. 11-18, Chicago, IL, 1996.

[3] P B. Dean, "Overview of Breast Cancer Screening," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp. 19-31, Chicago, IL, 1996.

[4] G. Cardenosa, "Mammography: An Overview," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp.3-10, Chicago, IL, 1996

[5] M. Sallam and K.W. Bowyer, "Registering Time Sequences of Mammograms Using Two-dimensional Image Unwarping Technique," *Proceedings of the 2nd International Workshop on Digital Mammography*, pp. 121-130, York England, 1994.

[6] N. Vujovic, P. Bakic, and D. Brzakovic, " Detection of Potentially Cancerous Signs by Mammogram Follow-up," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp. 421-424, Chicago, IL, 1996

[7] N. Vujovic and Brzakovic, " Establishing the Correspondence Between Control Points in Pairs of Mammographic Images," IEEE Trans. on Image Processing, vol. **6**, pp. 1388-1399, 1997.

[8] W. K. Zouras, et. al., "Investigation of a Temporal Subtraction Scheme for Computerized Detection of Breast Masses in Mammograms," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp. 411-415, Chicago, IL, 1996.

[9] M. A. Wirth and C. Choi, " Multimodal Registration of Anatomical Medial Images," *Australian Pattern Recognition Society, Conference on Image processing*, Oct. 1996.

[10] R.M. Ford, et. al., " Metrics for Scene Change Detection in Digital Video Sequences," *Multimedia Computing and Sys. 97 Proceedings IEEE International Conference on*, pp. 610-611, 1997.

[11] X. Dai, S. Khorram, "The Effects of Image Misregistration on the Accuracy of Remotely Sensed Change Detection," IEEE Trans. Geosci. Remote Sensing, vol. **36**, pp1566-1577,1998.

[12] P. Dhawan, et. al., "Iterative Principal Axes Registration Method for Analysis of MR-PET Brain Images," IEEE Trans. Biomed. Eng., vol. **42**, pp. 1079-1087,1995.

[13] C. L. Lin, Q. Zheng, R. Chellappa, L. S. Davis, X. Zhang," Site model supported monitoring of aerial images," *Computer Vision and Pattern Recognition*, June pp. 694-700,1994.

[14] L. Ott, *An Introduction to Statistical Methods and Data Analysis*, pp. 220-229, Wadsworth Pub. Co.,Belmont, Ca., 1977.

[15] R. C. Jain, et. al., "On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes," IEEE Trans. Patt. Anal. Mach. Intell., vol. **1**, pp. 206-213, 1979.

[16] R.L. Campbell, "Image Enhancement via morphological Filtering," *International Conference on Signal Processing Applications & Technology*, vol. **2**, pp. 1133-1137, Boston Ma. 1996.

[17] Z. Liang, "Tissue Classification and Segmentation of MR Images," IEEE Eng. in Medicine and Biology, pp. 81-85. Mar. 1993.

[18] Y. Wang, "MRI statistics and model-based MR image analysis," Ph.D. report, University of Maryland Graduate School, Baltimore, MD, April 1995.

[19] T. Lei, W. Sewchand, "Statistical Approach to X-Ray CT Imaging and Its Applications in Image Analysis — Part II: A New Stochastic Model-Based Image Segmentation Technique for X-Ray CT Image," IEEE Trans. Med. Imaging, vol.11, no. 1, pp. 62-69, Mar. 1992.

[20] Y. Wang, T. Adali, S.B. Lo, "Automatic Threshold Selection Using Histogram Quantization," J. Biomedical Optics, Vol. 2, No. 2, pp. 211-217, April 1997.

[21] C. A. Bouman, M. Shapiro, "A Multiscale Random Field Model for Bayesian Image Segmentation," IEEE Trans. Image Proc., Vol.3, No. 2, pp. 162-176, Mar. 1994.

[22] J. Zhang, J.W. Modestino, D. A. Langan, "Maximum-Likelihood Parameter Estimation for Unsupervised Stochastic Model-Based Image Segmentation," IEEE Trans. Image Proc., Vol.3 No. 4, pp. 404-420, July 1994.

[23] D.A. Langan, J.W. Modestino, J. Zhang, "Cluster Validation for Unsupervised Stochastic Model-Based Image Segmentation," IEEE Trans. Image Proc., Vol. 7 No. 2, pp. 180-195, Feb. 1998.

[24] Z. Liang, J.R. MacFall, D. P. Harrington, "Parameter Estimation and Tissue Segmentation from Multispectral MR Images," IEEE Trans. Medical Imaging, Vol. 13, No. 3, pp. 441-449, Sept. 1994.

[25] Dempster, A. P., Laird, N.M. and Rubin, D. B., "Maximum Likelihood from Incomplete Data via the EM algorithm," J. Roy. Soc. Statist., B., No. 1, pp. 1-38, 1977.

[26] Y. Wang, T. Adali, M.T. Freedman, and S. K. Mun, " MR Brain Image Analysis by Distribution Learning and Realization Labeling," Proc. 15th South. Biomed. Eng. Conf., pp. 133-136, Dayton Ohio, Mar. 1996.

[27] Y. Wang, T. Adali, S-Y Kung, and Z. Szabo, "... A Probabilistic Neural Network Approach," IEEE Trans. Image Proc., Vol. 7, No. 8, pp.1165-1181, Aug. 1998.

[28] Y. Wang, T. Lei, " A New Look at Finite Mixture Models in Medical Image Analysis," ISSIPNN, 1994, pp. 33-35.

[29] S. M. LaValle, S. A. Hutchinson, "A Bayesian Segmentation Methodology for Parametric Image Models," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 17, No. 2, Feb. 1995, pp 211-217.

[30] M. Cheriet, J.N. Said, C. Y. Suen,"A Recursive Thresholding Technique for Image Segmentation," IEEE Trans. Image Processing, Vol. 7 No. 6, June 1998, pp. 918-921.

[31] S. L. Sclove, " Application of the Conditional Population-Mixture Model to Image Segmentation," IEEE Trans. PAMI, Vol. PAMI-5, No. 4, July 1983, pp. 428-433.

[32] J. Zhang, J. W. Modestino, "A Model-Fitting Approach to Cluster Validation with Application to Stochastic Model-Based Image Segmentation," IEEE Trans. PAMI, Vol. 12, No. 10, Oct. 1990, pp. 1009-1017.

[33] Y. Delignon, A. Marzouki, W. Pieczynski,"Estimation of Generalized Mixtures and Its Application in Image Segmentation," IEEE Trans. Image Processing Vol. 6, No. 10, Oct. 1997, pp. 1364-1375.

[34] S. S. Saquib, C. A. Bouman, K. Sauer,"ML Parameter Estimation for Markov Random Fields with Application to Bayesian Tomography," IEEE Trans. Image Processing, Vol. 7, No. 7, July 1998, pp.1029-1044.

[35] C. Bouman, B. Liu,"Multiple Resolution Segmentation of Textured Images," IEEE Trans. PAMI, Vol. 13, No.2, Feb. 1991, pp. 99-113.

[36] K. Held et al., "Markov Random Field Segmentation of Brain MR Images," IEEE Trans. Medical Imaging, Vol.16, No. 6, Dec. 1997, pp. 878-886.

[37] T.M Chang, Y.H Liu, C.H. Chen, et al.,"Intermodality Registration and Fusion of Liver Images for Medical Diagnosis," Intelligent Information Systems 1997 ISS 1997 Proceedings.

[38] T.D. Zuk, M.S. Atkins, "A Comparison of Manual and Automatic Methods for Registering Scans of the Head," IEEE Trans. Medical Imaging, Vol. 15, No. 5, Oct. 1996 pp 732-744.

[39] WM. Wells III, P. Viola, H. Atsumi, et al.,"Multi-modal volume registration by maximization of mutual information," Medical Image Analysis Vol. 1, No. 1, pp35-51.

[40] A. Moskalik, P.L. carson, C.R. Meyer, J.B. Fowlkes, J.M. Rubin, et al.,"Registration of Three-Dimensional Compound Ultrasound Scans of the Breast for Refraction and Motion correction," Ultrasound Med. & Biol., Vol. 21 No. 6. pp 769-778, 1995.

[41] P. A. Van den Elsen, J.B. Antoine Maintz, et al.,"Automatic Registration of CT and MR Brain Images Using correlation of Geometrical Features," IEEE Trans. Medical Imaging Vol. 14 No. 2, June 1995, pp. 384-396.

[42] C. R. Maurer Jr., G. B. Aboutanos, et al."Registration of 3-D Images Using Weighted Geometrical Features," IEEE Trans. Medical Imaging, Vol. 15, Dec. 1996, pp. 836-849.

[43] P. A. Vand den Elsen, E. D. Pol, Max A. Viergever," Medical Image Matching – A review with Classification," IEEE Eng. Medicine and Biology, Mar. 1993, pp. 26-38.

[44] C.R. Maurer Jr., J.M. Fitzpatrick, et al.,"Registration of Head Volume Images Using Implantable Fiducial Markers," IEEE. Tra s. Medical Imaging Vol. 16 No. 4, Aug. 1997, pp. 447-462.

[45] C. Davatzikos, J. L. Prince and R.N. Bryan, " Image Registration Based on Boundary Mapping," IEEE Trans. Medical Imaging, Vol. 15, No. 1 Feb. 1996, pp. 112-115.

[46] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality Image Registration by Maximization of Mutual Information," IEEE Trans. Medical Imaging Vol. 16, No. 2, April 1997, pp. 187-198.

[47] R. J. Althof, M. G. J. Wind, J. T. Dobbins," A Rapid and Automatic Image Registration Algorithm with Subpixel Accuracy," IEEE Trans. Medical Imaging Vol. 16. No. 3, June 1997, pp. 308-316.

[48] C. R. Maurer, Jr., et al., " Registration of Head CT Images to Physical Space Using a Weighted Combination of Points and Surfaces," IEEE Trans. Medical Imaging , Vol. 17. No 5. Oct. 1998, pp. 753-761.

[49] C.A. Pelizzari et al.," Comparison of Two Methods for 3D Registration of PET and MRI Images," AIC of IEEE Eng. in Medicine and Biology Society, Vol. 13, No. 1, 1991 pp. 221-223.

[50] L. K. Arata, A. P. Dhawan," Iterative Principal Axes Registration: A New Algorithm for Retrospective Correlation of MR-PET Brain Images," AIC of IEEE Eng. in Medicine and Biology Society, Vol. 14, No. 7, 1992 pp. 2776-2778.

[51] S. C. Strother et al., " Quantitative Comparisons of Image Registration Techniques Based on High-Resolution MRI of the Brain," Journal of Computer Assisted Tomography, Vol. 18, No. 6, Nov/Dec. 1994, pp. 954-962.

[52] M. S. Brown et al.," Method for Segmenting Chest CT Image Data Using an Anatomical Model: Preliminary Results," IEEE Trans. on Medical Imaging Vol. 16 No. 6 Dec. 1997, pp. 828-839.

[53] L. K. Arata et al., "Three-Dimensional Anatomical Model-Based Segmentation of MR Brain Images Through principal Axes Registration," IEEE Trans. on Biomedical Engineering, Vol. 42. No. 11, Nov. 1995, pp. 1069-1078.

[54] K. Woods, J. Wang, M. T. Freedman, " Unsupervised Tissue Quantification and Segmentation from 3-D MRI Brain Images", IASTED SIP 1998. pp. 772-775.

[55] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, A.C. Evans : "Design and Construction of a Realistic Digital Brain Phantom" IEEE Trans. on Medical Imaging, Vol.17, No.3, p.463–468, June 1998.

[56] R.K.-S. Kwan, A.C. Evans, G.B. Pike : "An Extensible MRI Simulator for Post-Processing Evaluation" Visualization in Biomedical Computing (VBC'96). Lecture Notes in Computer Science, Vol. 1131. Springer-Verlag,1996. 135-140.

[57] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, A.C. Evans : "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database" NeuroImage, Vol.5, No.4, part 2/4, S425, 1997 – Proceedings of 3rd International Conference on Functional Mapping of the Human Brain¡ Copenhagen, May 1997.

[58] N. Vujovic, "Registration of Time-Sequences of Random Textures with Application to Mammogram Follow-up," Ph.D. report, Lehigh University, May1997.

[59] N. S. Vujovic, et. al., "Analogic Algorithm for Point Pattern Matching with Application to Mammogram Followup," 4th Workshop on Cellular Neural Networks and App., June 24-28, 1998.

[60] D. Brzakovic, et. al., " Mammogram Analysis by Comparison with Previous Screenings," Proc. of the 2nd International Workshop on Digital Mammography, York, England, pp. 131-139,1994.

[61] N. Vujovic and D. Brzakovic," Feature Point Identification and Regional Registration in Sequences of Non-Structured Texture Images," Proc. International Conference on Image Proc., vol. 3, pp. 156-159, 1995.

[62] D. Brzakovic, et. al., " Early Detection of Cancerous Changes by Mammogram Comparison," Proc. SPIE Visual Communications and Image Processing 1994, pp. 1520-1531, Chicago, IL, 1994.

[63] M. Sallam and K. Bowyer, " Detecting Abnormal Densities in Mammograms by Comparison to Previous Screenings," Proc. of the 3rd International Workshop on Digital Mammography, Chicago, IL, pp. 417-420,1996.

[64] M. Sallam, et. al. ,"Screening Mammogram Images for Abnormalities Developing Over Time," Proc. IEEE Nuclear Science Symposium and med. Imaging Conference, pp. 1270-1272, 1992.

[65] M. Abdel-Mottaleb, et. al., " Locating the Boundary Between the Breast Skin Edge and Background in Digitized Mammograms," Proc. of the 3rd International Workshop on Digital Mammography, Chicago, IL, pp. 467-470, 1996.

[66] R. Chandrasekhar and Y. Attikiouzel, " A Simple Method for Automatically Locating the Nipple on Mammograms," IEEE Trans. Med. Imaging, vol. 16, no. 5, pp. 483-494, 1997.

[67] F. L. Bookstein, " Principal Warps: Thin-Plate Splines and the Decomposition of Deformations," PAMI vol. 11, no. 6, pp. 567-585, June 1989.

[68] M. A. Wirth, et. al., "Point to Point Registration of Non-Rigid Medical Images Using Local Elastic Transformation Methods," IEE. Image proc. App. July 14-17 1997.

[69] M. A. Wirth, et. al., "A Nonrigid-Body Approach to Matching Mammograms," 7th International conference on Image proc. App., pp 484-488, 1999.

[70] Goshtasby, " Registration of Images with Geometric Distortions," IEEE Trans. Geo. Remote Sensing, vol. 26, no. 1, pp. 60-64, Jan. 1988.

[71] G.J. Ettinger, W.E.L. Gunson, et al. "Automatic registration for Multiple Sclerosis Change Detection" Proceeding of the IEEE Workshop on biomedical Image Analysis, Seattle, WA. 1994.

[72] B. M. Hemminger, et. al., " Evaluation of Digital Processing Methods for the Display of Digital Mammography," SPIE Conf. Image Display, SPIE vol. 3658, pp. 382-393, Feb. 1999.

[73] L. Brown, "A survey of image registration techniques", ACM Computing-Surveys, vol. 24. pp325-376, New York, 1992.

[74] R. Collins, et al., " Model matching and extension for automated 3D site modeling," Proc. ARPA Image Understanding Workshop, Washington D.C., pp. 197-204, April 1993.

[75] R. Collins, A. Hanson, E. Riseman, "Site model acquisition under the UMass RADIUS Project," Proc. ARPA Image Understanding Workshop, Monterey, CA., pp. 351-358 Nov. 1994.

[76] H.L. Van Trees," Detection, Estimation, and Modulation Theory," John Wiley and Sons, New York, 1968

[77] S. J. Orfanidis,"Optimum Signal Processing an Introduction," Mc Graw-Hill Pub. Company, New York, 1988.

[78] Q. Zheng, R. Chellappa," A computational vision approach to image registration." IEEE Trans on Image Processing, vol. 2, pp. 311-326, 1993.

[79] R. Chellappa, et. al., "Site-Model-Based Monitoring of Aerial Images," Proc. ARPA Image Understanding Workshop, pp. 295-318, 1994.

[80] X. Zhang, et. al.,"Automatic Image to Site Model Registration," Proc. Intl. Conf. on Acoustics, Speech, and Signal Processing, pp. 2164-2167, Atlanta, GA, May 1996.

[81] H. Yildirim, et. al., "Temporal Change Detection by Principal Component Transformation," IEEE International Geoscience and Remote Sensing Symposium, vol. 2, pp. 1227-1229, 1995.

[82] L. Bruzzone, S. B. Serpico, "An Iterative Technique for the Detection of Land-Cover Transitions in Multitemporal Remote-Sensing Images," IEEE Tans. on Geoscience and Remote sensing, Vol. 35, no.4, pp. 858-866. 1997.

[83] T. Yamamoto, et. al., " A Change Detection Method for Remotely Sensed Multi-spectral and Multi-Temporal Images using 3-D Segmentation," IEEE International Geoscience and Remote Sensing Symposium vol. 1, pp 77-79, 1999.

[84] J.B. Antoine Maintz ,M. A. Viergever, "A survey of medical image registration," Medical Image Analysis, vol.2, pp. 1 – 36, 1998.

[85] A. K. Jain, "Fundamentals of Digital Image Processing," Prentice-Hall, 1989.

[86] B. Kolman, "Introductory linear algebra with applications," Macmillan Publishing Co., 1988.

[87] K. I. Laws, "Rapid Texture Identification," Proc. SPICE Conf. Image Processing for Missile Guidance, pp. 376-380, 1980.

[88] W. F. Good, et. al., "Image modification for display of temporal sequences of mammograms," Medical Imaging 2000: Image Display and Visualization, Proceedings of SPIE, Vol. 3976, pp. 174-184, 2000.

Patient site model supported change detection

K. Woods[a], M. McClain[a], Y. Wang[a], and M. T. Freedman[b]

[a]Department of Electrical Engineering and Computer Science
The Catholic University of America, Washington, DC 20064, USA
[b]Department of Radiology and the Lombardi Cancer Center
Georgetown University Medical Center, Washington, DC 20007, USA

ABSTRACT

This paper reports the development of a non-rigid registration technique to bring into alignment a sequence of a patient's single-view mammograms acquired at different times. This technique is applied in a patient site model supported change detection algorithm with a clinical goal of lesion detection and tracking. The algorithm flow contains four steps: preprocessing, image alignment, change detection, and site model updating. The preprocessing step includes segmentation, using standard finite normal mixture and markov random field models, morphological processing, monotony operators, and guassian filtering. The site model in this research is composed of object boundaries, previous change, potential control points, and raw/segmented images. In the alignment step, the current mammogram is aligned to the site model using a two step process consisting of principle axis of the skin line followed by thin-plate spline using matched points from the potential control point pool. With the assumption of minimal global change, subtraction and thresholding will be used to create the change map that highlights significant changes. Finally, the change information will be used to update the site model. This two-step registration process facilitates change detection by aligning corresponding regions of mammograms so local change analysis can be performed in a coherent manner. The result of the change detection algorithm will be a local change and a patient specific site model showing past and present conditions.

Keywords: Computer aided diagnosis, Change detection, Principle axis registration, Mutual information registration, Segmentation, Site model, Feature extraction

1. INTRODUCTION

Breast cancer is one of the leading causes of death among women today. To combat this problem doctors use medical imaging as a mechanism to determine if any additional tests should be performed. For instance, the mammography has proven to be the only way to detect cancer at its earliest stages, thus improving the patient survival probability. This type of study is called breast cancer screening and usually is limited to asymptomatic women where caniocaudal (CC) and mediatorial oblique (MLO) mammographic views are analyzed.[2] Tumor size has an apparent relationship to tumor grade at the time of diagnosis. So, starting at approximately age 40, most women have screening mammograms performed periodically in effort to detect the existence or onset of a cancerous condition in the breast. These images are usually reviewed manually by a radiologist who views a two mammogram sequence composed of a single view, of a single breast, acquired at different times, looking for visually apparent change between the mammograms.[2] Studies have shown a correspondence between tissue change and underlying biological change. This change is important for applications such as treatment monitoring and lesion diagnosis. The review of this massive volume of data by the radiologist results in missed tumors, delayed detection and false positives which ultimately cause a reduced life expectation upon detection, unnecessary patient call backs, and unneeded biopsies.

To reduce some of the load on the radiologist and to improve diagnosis accuracy, development of automated approaches have been considered,[3] [4] using a single view of one breast and[6] using single view multiple (left and right) breasts. Use of multiple breast leads to additional problems because women typically have significantly different structures between left and right breasts. This causes natural asymmetry to be flagged as change.[3] The single breast approaches, on the other hand, do not have the problem of dealing with asymmetry. Generally, single breast approaches contain three main steps: (1) preprocessing of the images searching for control points or regions for use in registration, (2) registration, to align the images into a common framework, and (3) detection and analysis of

In *Medical Imaging 2000: Image Processing*, Kenneth M. Hanson, Editor,
Proceedings of SPIE Vol. 3979 (2000) • 1605-7422/00/$15.00

1095

local change. The preprocessing is generally handled by classical image processing techniques such as segmentation, morphological filtering, edge detection, and feature extraction.

The main problem with automating change detection analysis is performing automatic image registration between mammogram images. This difficulty is attributed to three main problems. First mammograms are complex images that do not contain any clearly defined landmarks. Secondly, differences in breast positioning and compression during acquisition could cause two visually different images. Finally, breast sizes and consistency can vary with time (e.g. weight loss and surgery).

The group[3] approached these problems by extracting the dense tissue of the breast from both images using segmentation, and then performing a sequence of two thin-plate spline (TPS) registrations.[7] The first TPS uses control points extracted from the smoothed dense tissue boundary. These control points are obtained by determining the points of maximum curvature on the boundary of both images and comparing statics of surrounding intensities to determine the correspondence. The second TPS uses control points extracted from the dense tissue itself. Correspondence between points is performed by a signature match between images which then feeds an accumulator matrix.[3] This approach has problems when the dense tissue does not occupy a large percentage of the image which typically occurs in radio-lucent breast.[1] In cases like this, error occurs in transforms when the point to be transformed is far away from the control points thus reducing the effect of the control points.

[4] considers these same problems by asserting that accurate registration of mammograms is intractable except with elastic transforms, and the only solution is regional registration.[5] In regional registration localized areas of the two mammograms are aligned based on their distance from control points. In their approach, monotony operators are used to extract vertical and horizontal elongated structures which they assume to be stable between images in the sequence. These structures correspond to blood vessels and ducts. A three-pass Gaussian filter is used on the original mammogram to mask less prominent structures. This reduces the complexity and limits the monotony operators to detecting the dominate structures. The cross points of these horizontal and vertical structures make up the pool of potential control points. Correspondence between the current image control points and reference image control points is accomplished by comparing the respective signatures. To localize the area where signatures are compared, the nipple location in both images are used to determine a neighborhood region. This reduces processing and decreases the probability of false alarm. These values are then passed into a thresholded accumulator matrix for final point selection. Using these control points, regions (of any shape) are determined on the current image by calculating the distance from a subset of control points. This method over comes the erroneous interpolation problem experienced by,[3] but the algorithm uses ad hoc point matching criteria, window size selection, and threshold determination. In addition,[5] assumes a small misregistration that restricts the generality of this approach. Both[3] and[4] mainly address registration so, simple change detection methodologies based on difference image analysis and wavelets respectively.

To address the following problems: control point correspondence issues, TPS interpolation problems, link between registration and change detection, and restrictive assumptions faced by current mammogram registration algorithms, we propose a multi-step registration algorithm that aligns non-rigid objects to a common frame called a site model for change detection.[11] The site model is a mathematical model that over time describes the image scene (i.e. object locations etc.). This allows for the consideration of more than adjacent mammograms (in time) in the analysis which can improve detection probability by providing a complete history of previous conditions. The site model contains various types of data, called site model parameters, such as known anatomical structures, landmark points, previous tumor locations, segmented and raw data, suspected lesion locations and other informational notes. Site model parameters can be generated from preprocessing the raw reference image with segmentation, edge detection, feature extraction, or simple user input such as previous tumor locations, area of interest, and landmarks. The multi-step registration change detection algorithm has three main steps: initial registration, final registration and change detection analysis. Figure 1a shows a block diagram of the flow.

Initial registration considers all of the breast tissue as a solid object and performs principle axis registration (PAR) to correct for large misregistration between images. Final registration is performed by a polynomial based registration algorithm to handle non-rigid deformation that could occur between images. The affine polynomial is used to represent the mapping function. Polynomial based algorithms depend heavily on the existence of control points between the images. To obtain the control points, we follow a modified version of the approach discussed in.[5] The approach is modified by using the Pearson correlation coefficient[12] to match the two potential control point signatures. The change analysis is performed with a difference image, histogram, and visual inspection.
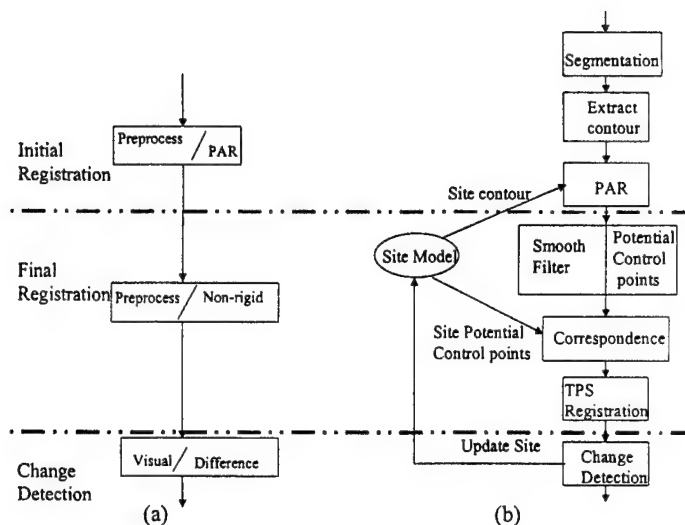
Figure 1. (a) Algorithm flow. (b) Detailed algorithm flow.

During the development of this algorithm, several assumptions were made in order to bound the scope of this paper. First, the mammograms are assumed to be CC and MLO views only (i.e. screening mammograms) of the same patient acquired overtime. Second, the radiologist initializes the site model parameters by identifying an area of interest window and other prominent landmark points in the first image of the sequence. Change is then calculated for the pixels in this windows. Third, the type of change was limited to growth of a mass, or shrinkage of a mass. Microcalcifications changes will not be addressed in this paper. Fourth, the amount of misregistration is found to approximately be $+/-25$ degrees rotation and translation between the reference and current mammograms.

The contributions of this paper are as follows: the introduction of a multi-step registration algorithm consisting of a rigid first step (PAR) followed by a non-rigid second step (TPS/global affine), verification of[4] control point methodology, improvement of[5] signature match algorithm using correlation coefficient, and finally introduction of the site model concept to medical imaging which enables analysis of the results of more than two mammograms through site model build up. The paper is organized in five sections. Section I introduces the topic. Section II describes the materials and methods used while section III presents the simulation results. Finally, sections IV and V contain discussion and conclusion respectively.

2. MATERIAL AND METHODS

The processing algorithm is broken into three phases: initial registration, final registration, and change detection. The complete algorithm flow is shown in Figure 1b. The algorithm starts by performing PAR to obtain an initial registration. Next, the initial registration is fine tuned using a global affine based registration process. To finish the processing, change detection using a difference based analysis is done. In this section, the theory of each phase is discussed in detail.

2.1. Initial Registration

The initialization of the transformation process is performed by preprocessing the images. Preprocessing consists of image segmentation and morphological filtering. A statistical based algorithm is used for the segmentation process. The algorithm models image intensity distributions of a NxN image with a standard finite normal mixture model with three degrees of freedom as shown below.

$$f(x/\varphi) = \prod_{i=1}^{N^2} \sum_{k=1}^{K} \pi_k g(x/\theta_k) \tag{1}$$

$$\theta_k = \mu_k, \sigma^2 \qquad ; \qquad \varphi_k = \pi_k, \theta_k \tag{2}$$

where the unknown parameters are mean μ_k, variance σ_k^2, membership π_k, and K is the number of classes assumed to be in the image and g is a Gaussian kernel. Once these parameters are estimated using expectation maximization (EM) algorithm, pixels are labeled using Contextual Bayesian Relaxation CBRL which considers neighborhood relationships in pixel assignments.

After segmentation is complete, the skin line is extracted by grouping pixels that represent breast tissue into a single class. This operation forms a binary image which serves two purposes. First, the binary image serves as a mask that limits processing to the tissue regions of the image. Second, the binary image feeds a morphological filter designed to extract the breast contour.

Morphological filtering can be used to enhance binary images, by processing them using a kernel called a structuring element. The structuring element is designed to shape, distort, or filter the object in a specific manner. The two morphological operations considered in this research are dilation and erosion.

Dilation is explained by the following equations:

$$G = F \oplus H \tag{3}$$

Erosion is explained by the following equations:

$$G = F \ominus H \tag{4}$$

G is the processed image, F is the original image, H is the structuring element, $\oplus$ is Minkowski vector addition, $\ominus$ is Minkowski vector subtraction.

Erosion and dilation operations have the visual effect of thickening and thinning the contents of the image by processing it with the structuring element. To obtain the contour, the eroded image is subtracted from the dilated image and then the indices of the contour are ordered to yield the graphics representation of the object.

Given the graphical representation of the breast outline, PAR registration is performed between the contour stored in the site model (reference mammogram) and the float image (current mammogram). Principle axes registration is an intrinsic registration method that aligns bodies based on their moments of inertia. PAR assumes that the data sets being aligned have significant features in common.

Before the bodies are rotated and scaled, they are translated so that their centroids are collocated. The algorithm then computes the primary moments of inertia of the objects and the magnitudes of these moments through singular value decomposition (SVD). The moments of inertia are represented by eigenvectors and the magnitudes are the associated eigenvalues.

The algorithm steps are as follows:

1) Calculate the sample covariance matrix, C, of the data set.

$$C = \frac{1}{L} \sum_{i=1}^{L} (q_i)^T (q_i) \tag{5}$$

where q_i equals one of the L contour points with a center of gravity at the origin.

2) Use SVD to obtain the eigenvectors, V, and eigenvalues, the diagonal of Φ, of C.

$$CV = V\Phi \tag{6}$$

$$V^T CV = \Phi \tag{7}$$

3) Determine the scaling matrix for converting object a to the scale of object b.

$$\Phi_b S^2 = \Phi_a \tag{8}$$

where S is the scale factor.

4)Form the transformation matrix. The scaling and rotation matrix in one expression is given below.

$$U = V_A^T S V_B \tag{9}$$

5) Apply U to the complete image.

2.2. Final Registration

Similar to initial registration, final registration is divided into two portions, preprocessing and transformation. The goal of preprocessing here is to obtain control points between the image pair. This is achieved by first passing the PAR transformed image through a multiple pass Gaussian kernel to blur the fine details so only prominent structures are present in the image. The blurred image is then processed with two modified monotony operators: one for horizontal elongated structures and one for vertical elongated structures as presented in.[5] The monotony operators are defined by two overlapping rectangular neighborhoods, one small and one large, centered around a pixel (i,j). The operators work as follows: the pixel at (i,j) is labeled one if the number of pixels in the large neighborhood that are larger than g_{max}, exceeds a threshold τ. Otherwise, the operator assigns a zero to the pixel (i,j). g_{max} is defined as the maximum gray level in the small neighborhood surrounding the pixel (i,j). The vertical and horizontal operators are defined by the following relations

vertical:

$$a = \{(k,l)|k = 1, -p \le l \le p\} \tag{10}$$
$$A = \{(m,n)|m = 1, -q \le n \le q\}$$

horizontal:

$$a = \{(k,l)|l = 1, -p \le k \le p\} \tag{11}$$
$$A = \{(m,n)|n = 1, -q \le m \le q\}$$

$$q > p, \tau = (q - p) \tag{12}$$

where a is the small neighborhood of length p and A is the large neighborhood of length q. The potential control points are defined as the centroids of the crossing of vertical and horizontal elongated structures. This is implemented by applying a logical AND operation to the vertical elongated structures image Λ and horizontal elongated structures image Γ.

$$\Upsilon = \Gamma \odot \Lambda \tag{13}$$

Next, an attempt is made to match each of the potential control points in the site model (older image) with potential control points in the new image. A match is considered valid when a location criteria and signature match criteria are satisfied. The location criteria simplifies the search requirement by limiting the search area. The location criteria requires the potential point $O_q(x_q, y_q)$ in the new image to be contained by a kxl window centered around the point x_c. x_c is the intersection point between a circle centered around the estimated nipple location $O_n(x_n, y_n)$ and a straight line between O_n and potential control point O_q. The slope of the line is equal to the slope of
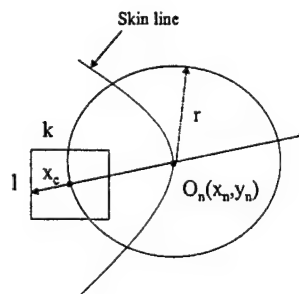
Figure 2. Matching window location on new mammogram

a similar line between the potential control point $O_p(x_p, y_p)$ in the site model (old image) and O_o the nipple location in the old image. The equations for the circle and line are shown below.

$$y = \frac{y_p - y_o}{x_p - x_o}(x - x_o) + y_n \tag{14}$$
$$(x - x_n)^2 + (y - y_n)^2 = (x_o - x_p)^2 + (y_p - y_o)^2$$

Figure 2 shows a pictorial representation of the window.

After passing the location criteria, the signatures of all potential control points contained, in the local window, are matched against the signature of O_p by calculating the Pearson correlation coefficient.[12]

$$\rho = \frac{SS_{xy}}{\sqrt[2]{SS_{xx}SS_{yy}}} \tag{15}$$
$$SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{h}$$

y is the h point signature of O_p.

The signatures are designed to capture the characteristics of the elongated structures. The signatures are formed by creating an elongated image by logical ORing the vertical and horizontal structure images to obtain Ω.

$$\Omega = \Gamma \oplus \Lambda \tag{16}$$

Then a mxn window is rotated $\Delta\xi$ steps around the control point and the number of nonzero pixels for each step is counted and stored as the signature. Because non-rigid deformation could occur between images the corresponding control point signature could be a circularly shifted version of the reference control point signature as seen in Figure 3. To consider this problem, the complete signature of the new image control point is circularly shifted by one sample and then Pearson matched. The highest Pearson between all shifts is taken to be the resulting Pearson value for that (O_p, O_q) pair.
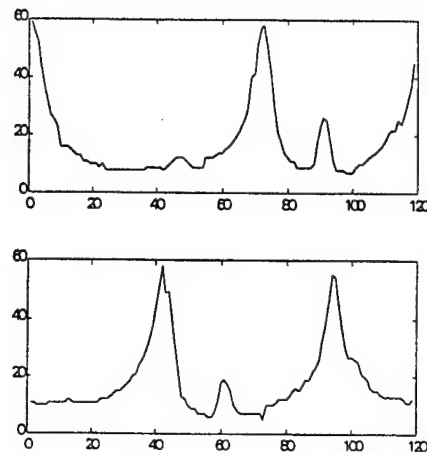
Figure 3. Potenial control point signature with corresponding shifted version
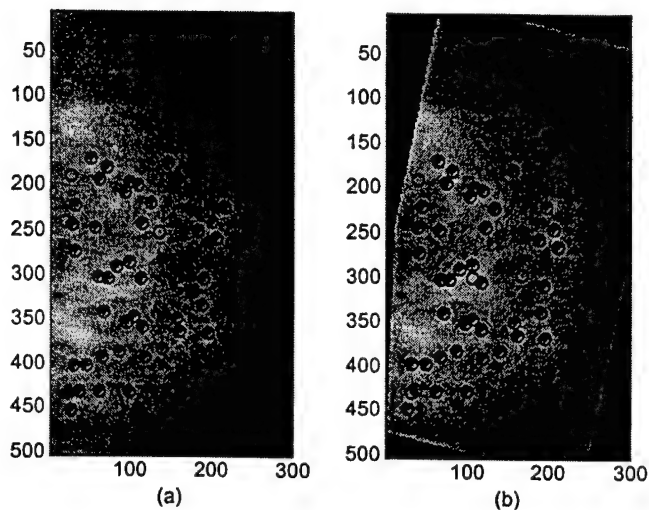

Figure 4. Potential and Matched Control points

The Pearson results for a (O_p, O_q) pair are stored in a modified accumulator matrix. The accumulator matrix is a $N_o x N_n$ matrix where N_o and N_n are the number of potential control points in the site model (old) and new images respectively. So each element is a (O_p, O_q) pair. In traditional accumulator formulations, the element (O_p, O_q) is incremented each time point O_p matches point O_q, but in this research we put the maximum Pearson correlation coefficient the element corresponding to (O_p, O_q). The final match is performed by taking the maximum value down the columns and zeroing the other column entries for that column. This is followed by taking the maximum value in each row and zeroing the other row entries. The resulting matrix should contain only one nonzero value per row and column. The nonzero elements are the control points as seen in Figure 4 where o and $*$ are the potential and real control points respectively .

The next step in the processing flow is performing the final registration. This registration is performed by a global

affine modified thin-plate spline transform that handles non-rigid deformations. Generally, the goal in registration is to obtain a transform T_A such that one of the images could be transformed into correspondence with the other. In general, an image mapping transform is represented by

$$T_A(x,y) = (f_x(x,y), f_y(x,y)) \tag{17}$$

where $f_x(x,y)$ is the mapping function for x coordinate of (x,y) and $f_y(x,y)$ is the mapping function for the y component of (x,y). The mapping function for affine are shown below.

$$f(x,y) = w_0 + w_1 x + w_2 y \tag{18}$$

In order to use $f(x,y)$ to transform the image the coefficients w_0, w_1, and w_2 must be estimated. Using the control points determined form the previous phase, we use a least square approach to determine the coefficients. The least squares formulation starts with following relation

$$(u,v) = (f_x(x,y), f_y(x,y)) \tag{19}$$

where (u,v) is the location of the point (x,y) in the new image. From this equation the error is derived

$$E = \sum_{i=1}^{n} \left[(u_i - f_x(x,y))^2 + (v_i - f_y(x,y))^2 \right] \tag{20}$$

The above equations lead to the normal equations

$$\sum_{i=0}^{m} \sum_{j=0}^{i} a_{ij} \left[\sum x_k^j y_k^{i-j} x_k^\beta y_k^{\alpha-\beta} \right] = \sum_{k=1}^{n} u_k x_k^\beta y_k^{\alpha-\beta} \tag{21}$$

where $\alpha = 0....m$ and $\beta = 0....\alpha$.

The coefficients for the y mapping functions are found in a similar fashion. Once both mapping functions are found, the images can be transformed pixel by pixel. Since we are considering registration of images of the same modality, gray level resmapling is optional. We choose to use a simple nearest neighborhood look up table to assign new labels of the transformed data.

2.3. Change Detection

The final phase of processing is change detection The change will be measured pixel-wise inside a local window that was identified by the radiologist. Visual inspection and difference image thresholding will be used to determine change. Difference image thresholding consists of taking the absolute value of the reference minus the current image bounded by a threshold. If the absolute value is larger than the threshold then the pixel is labeled changed, otherwise, the pixel is labeled unchanged. The formulation is show below.

$$D = abs(R_f - R_s)$$
$$imag(i,j) = \begin{Bmatrix} 1; D(i,j) > \gamma \\ 0; D(i,j) < \gamma \end{Bmatrix} \tag{22}$$

where D is the difference image, γ is the threshold, R_f is the transformed image, and R_s is the site model image.

2.4. Performance

To evaluate the performance of the algorithm, a phantom sequence and five mammogram sequences (two or more mammograms) were considered. The phantom sequence was created by perturbating a real mammogram with a simulated mass (block of constant value pixels that are Gaussian filtered) and misaligning it with the reference image. Using the phantom sequence the amount of change detected is calculated on a pixel bases. For the real mammograms change will be compared to the radiologist marked regions. The real mammograms were initially digitized at 100 microns yielding 2200x2400, but were down sampled to 500 microns 500x300 to make processing tractable.
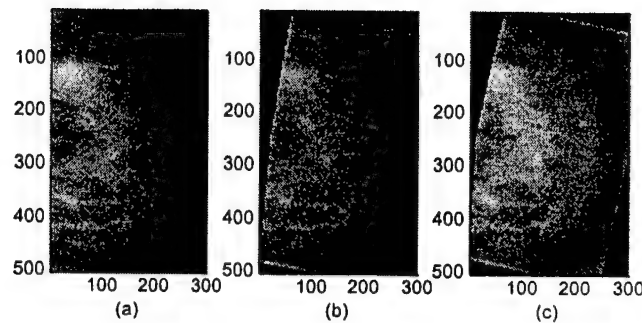
Figure 5. Sequence Registration (a) Site Model (b) first image in sequence (c) second image

2.5. Algorithm

The algorithm can be summarized into three main steps as outlined below.

Initial Registration

    Preprocess mammogram for skin line and tissue mask

    Use PAR on breast tissue as if it is a solid

Final Registration

    Preprocess the mammograms searching for control points and transform coefficients

    Use affine registration on the window of interest

Change Analysis

    Perform difference and threshold followed by visual inspection.

    Update change map located in the site model

3. RESULTS AND DISCUSSION

The simulation objective was to demonstrate the potential use of this algorithm in change detection in a mammogram sequence. Change detection relies heavily on the registration process; without registration there is no reasonable way to obtain correspondence between the images. As discussed previously, control point selection is a major step in the registration algorithm. We followed the formulation developed by,[5] but varied the approach with the use of Pearson correlation coefficient in signature matching. The improved signature matching criteria yielded, on average, more than double the control points as the matching criteria discusses in.[5] The additional control points improve registration, change detection, and site model build up. In Figure 5 we applied our multistep registration algorithm to the image phantom sequence where (a) was the site model image, (b) is registered version of first image, and (c) is a registered version of the last image in sequence. From this figure we see that general alignment was obtained which reduced the overall global change noise as shown in Figure 6, global change without registration and Figure 7 global change after registration. With the reduction of global change noise, the local change window now is dominated by change caused by mass growth or shrinkage. Histogram comparison difference and visual inspection where used to determine change. Figure 8 shows the intensity histogram of two region identified as change. Basic difference analysis was attempted, but was insufficient to deal with the more complex intensity patterns. Change detection not only highlights existence of possible changed regions, but when combined with the site model provides
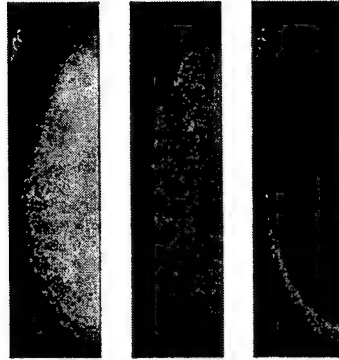
Figure 6. Difference image unregistered pair
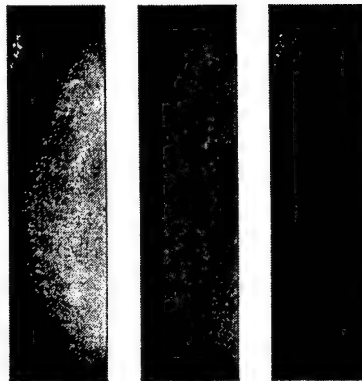
Figure 7. Difference image registered pair
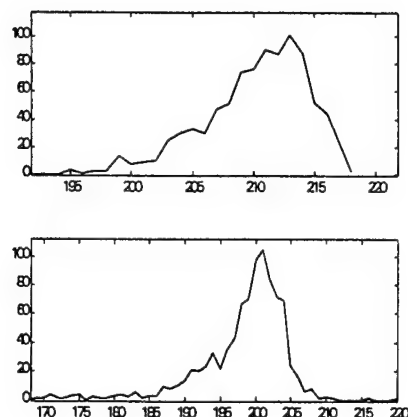
Figure 8. Histogram comparison between the two local change windows
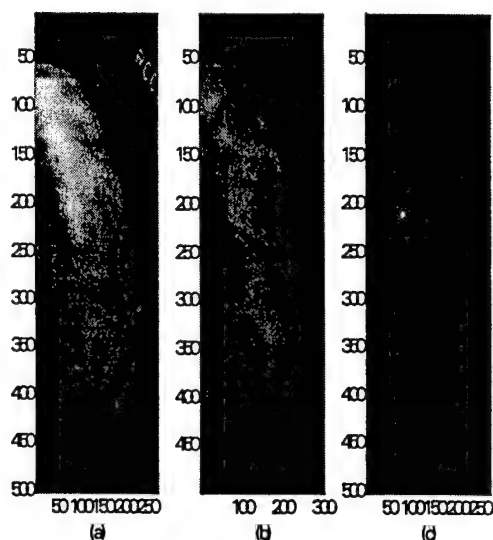
Figure 9. Change detection process (a) site image (b) registered image (c) detected change

a patient history by showing site progression. For the phantom sequence, the algorithm detected 19 of the 100 changed pixels in the first image and 207 of the 400 changed pixels in the second image. This shows a mass growth of about 188 pixels between the last two images in the sequence. This is not exact detection, but would be sufficient to flag a radiologist to review the area. Figure 9 shows the algorithm applied to a real mammogram sequence where the detected change is shown in (c). The main results of this study consisted of the automatic alignment of mammograms, detection of change in a local window, and implementation of a mechanism to store and build up patient information via the site model.

4. CONCLUSION

This paper considered the development of an multistep algorithm to perform change detection with the added benefit of site memory build up (site model). The algorithm consists of three main steps: initial registration, final registration, and change detection. Each of these processes interacts with the site model. The initial registration uses PAR of the complete breast with the principle axes of the site model image. The final registration uses the global affine derived from parameters in the site model. The global affine is used to handle non-uniform deformation between image sets. From the results, we have shown that the float image is aligned enough to match a local windows surround the area of interest to perform change detection. Occasional change was missed because of inaccuracies in the registration process, but this is overcome by selecting a layer analysis window. Currently, we are studying methods to make the registration more robust and image independent by selecting more control points and stronger methods to verify control point correspondence. We believe this should lead to better change detection with the ability to quantify actual change size.

5. ACKNOWLEDGMENT

REFERENCES

1. D. B. Kopans, *Breast Imaging*, J.B. Lippincott Co.,Philadelphia, Pa., 1989.
2. G. Cardenosa, "Mammography: An Overview," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp.3-10, Chicago, IL, 1996.
3. M. Sallam and K.W. Bowyer, "Registering Time Sequences of Mammograms Using Two-dimensional Image Unwarping Technique," *Proceedings of the 2nd International Workshop on Digital Mammography*, pp. 121-130, York England, 1994.
4. N. Vujovic, P. Bakic, and D. Brzakovic, " Detection of Potentially Cancerous Signs by Mammogram Follow-up," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp. 421-424, Chicago, IL, 1996
5. N. Vujovic and Brzakovic, " Establishing the Correspondence Between Control Points in Pairs of Mammographic Images," IEEE Trans. on Image Processing, vol. **6**, pp. 1388-1399, 1997.
6. W. K. Zouras, et. al., "Investigation of a Temporal Subtraction Scheme for Computerized Detection of Breast Masses in Mammograms," *Proceedings of the 3rd International Workshop on Digital Mammography*, pp. 411-415, Chicago, IL, 1996.
7. M. A. Wirth and C. Choi, " Multimodal Registration of Anatomical Medial Images," *Australian Pattern Recognition Society, Conference on Image rocessing*, Oct. 1996.
8. R.M. Ford, et. al., " Metrics for Scene Change Detection in Digital Video Sequences," *Multimedia Computing and Sys. 97 Proceedings IEEE International Conference on*, pp. 610-611, 1997.
9. X. Dai, S. Khorram, "The Effects of Image Misregistration on the Accuracy of Remotely Sensed Change Detection," IEEE Trans. Geosci. Remote Sensing, vol. **36**, pp1566-1577,1998.
10. P. Dhawan, et. al., "Iterative Principal Axes Registration Method for Analysis of MR-PET Brain Images," IEEE Trans. Biomed. Eng., vol. **42**, pp. 1079-1087,1995.
11. C. L. Lin, Q. Zheng, R. Chellappa, L. S. Davis, X. Zhang," Site model supported monitoring of aerial images," *Computer Vision and Pattern Recognition*, June pp. 694-700,1994.
12. L. Ott, *An Introduction to Statistical Methods and Data Analysis*, pp. 220-229, Wadsworth Pub. Co.,Belmont, Ca., 1977.
13. R. C. Jain, et. al., "On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes," IEEE Trans. Patt. Anal. Mach. Intell., vol. **1**, pp. 206-213, 1979.
14. R.L. Campbell, "Image Enhancement via morphological Filtering," *International Conference on Signal Processing Applications & Technology*, vol. **2**, pp. 1133-1137, Boston Ma. 1996.

Computerized Radiographic Mass Detection – Part I:

Lesion Site Selection by Morphological Enhancement and Contextual Segmentation

Huai Li[1,3], Yue Wang[2,3], K. J. Ray Liu[1]
Shih-Chung B. Lo[3], and Matthew T. Freedman[3]

[1]Electrical Engineering Department and Institute for Systems Research
University of Maryland at College Park, College Park, Maryland 20742

[2]Department of Electrical Engineering and Computer Science
The Catholic University of America, Washington, DC 20064

[3]Department of Radiology, Georgetown University Medical Center
Washington, D.C. 20007
huaili@eng.umd.edu and kjrliu@eng.umd.edu

ABSTRACT

This paper presents a statistical model supported approach for enhanced segmentation and extraction of suspicious mass areas from mammographic images. With an appropriate statistical description of various discriminate characteristics of both true and false candidates from the localized areas, an improved mass detection may be achieved in computer-aided diagnosis. In this study, one type of morphological operation is derived to enhance disease patterns of suspected masses by cleaning up unrelated background clutters, and a model-based image segmentation is performed to localize the suspected mass areas using stochastic relaxation labeling scheme. We discuss the importance of model selection when a finite generalized Gaussian mixture is employed, and use the information theoretic criteria to determine the optimal model structure and parameters. Examples are presented to show the effectiveness of the proposed methods on mass lesion enhancement and segmentation when applied to mammographical images. Experimental results demonstrate that the proposed method achieves a very satisfactory performance as a pre-processing procedure for mass detection in computer-aided diagnosis.

Keywords: Morphological filtering, finite mixture, relaxation labeling, information criterion, image enhancement, image segmentation.

I Introduction

In recent years, several computer-aided diagnosis (CAD) schemes for mass detection and classification have been developed [1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13]. Though it may be difficult to compare the relative performance of these methods, because the reported performance strongly depends on the degree of subtlety of masses in the selected database, accurate selection of suspected masses is considered a critical and first step due to the variability of normal breast tissue and the lower contrast and ill-defined margins of masses [3, 6], and since no subtle masses should be missed before any further analysis.

A number of image processing techniques have been proposed to perform suspicious mass site selection. Kobatake *et al* [1] proposed using a iris filter to detect tumors as suspicious regions with very weak contrast to their background. Sameti *et al* [7] used fuzzy sets to partition the mammographic image data. Lau and Yin *et al* independently proposed using bilateral-subtraction to determine possible mass locations [9, 13]. Some other investigators proposed using pixel-based feature segmentation of spiculated masses [4, 8]. Kegelmeyer has reported promising results for detecting spiculated tumors based on local edge characteristics and Laws texture features [8]. Karssemeijer *et al* [4] proposed to identify stellate distortions by using the orientation map of line-like structures. Recently, Petrick *et al* [6] proposed a two-stage adaptive density-weighted contrast enhancement filtering technique along with edge detection and morphological feature classification for automatic segmentation of potential masses. Kupinski and Giger [3] presented a radial gradient index-based algorithm and a probabilistic algorithm for seeded lesion segmentation.

Nevertheless, to our best knowledge, few work has been dedicated to improve the task of lesion site selection although it is indeed a very crucial step in CAD. Especially, few studies have used and justified model-based image processing techniques for unsupervised lesion site selection [11]. Zwiggelaar *et al* developed a statistical model to describe and detect the abnormal pattern of linear structures of spiculated lesions [2]. In their work, the probability density function of the observation vectors for each class is assumed to be normal. we have experienced that the "normal" distribution for each class is nor true. Li *et al* proposed using a Markov random field model to extract suspicious masses for mass detection [11]. In their study, most

of model parameters were chosen empirically, and the mammogram was segmented into three regions (background, fat, and parenchymal or tumors).

Stochastic model-based image segmentation is a technique for partitioning an image into distinctive meaningful regions based on the statistical properties of both gray-level and context images. A good segmentation result would depend on suitable model selection for a specific image modality [16, 17] where model selection refers to the determination of both the number of image regions and the local statistical distributions of each region. Furthermore, a segmentation result would be improved with pre-enhanced pattern of interest being segmented. The only assumption for suspected mass site selection is that suspected mass areas should be brighter than the surrounding breast tissues which is valid for most of the real cases. When some masses lie either within an inhomogeneous pattern of fibroglandular tissue or are partially or completely surrounded by fibroglandular tissue, enhancement of mass-related signals is important.

Fig. 1 shows a general block diagram of CAD systems. The Part I of this paper focuses on "Image Processing" block, to just automatically pick up all possible lesion sites. We aim on two essential issues in the stochastic model-based image segmentation: enhancement and model selection. Based on the differential geometric characteristics of masses against the background tissues, we propose one type of morphological operation to enhance the mass patterns on mammograms. Then we employ a finite generalized Gaussian mixture (FGGM) distribution to model the histogram of the mammograms where the statistical properties of the pixel images are largely unknown and are to be incorporated. We incorporate the EM algorithm with two information theoretic criteria to determine the optimal number of image regions and the kernel shape in the FGGM model. Finally, we apply a contextual Bayesian relaxation labeling (CBRL) technique to perform the selection of suspected masses. The major differences of our work from the previous work [1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13] are that:

1. We have presents a new algorithm of morphological filtering for image enhancement in which the combined operations are applied to the original gray tone image and the higher sensitive lesion site selection of the enhanced images are observed.

2. We have justified and pilot tested the finite generalized Gaussian mixture (FGGM) dis-

tribution in modeling mammographic pixel images together with a model selection procedure based on the two information theoretic criteria. This allows an automatic identification of both the number (K) and kernel shape (α) of the distributions of tissue types.

3. We have developed a new algorithm (CBRL) for segmenting mass areas where the comparable results are achieved as those using Markov random field model based approaches while with much less computational complexity.

The presentation of this paper is organized as follows. In Section II, the proposed dual morphological operation enhancement technique is described in detail. The theory and algorithm on FGGM modeling, model selection, and parameter estimation are presented in Section III. This is followed by a discussion on the selection of suspicious masses using the CBRL approach. Evaluation results are given and discussed in Section IV. Finally, the paper is concluded by Section V.

II Morphological Enhancement

One of the main difficulties in suspicious mass segmentation is that mammographic masses are often overlapped with dense breast tissues. Therefore, it is necessary to remove bright background caused by dense breast tissues while preserving the features and patterns related to the masses. For this purpose, background correction is an important step for mass segmentation. We propose a mass pattern-dependent background removal approach using morphological operations.

II.1 Morphological Filtering Theory

Morphological operations can be employed for many image processing purposes, including edge detection, region segmentation, and image enhancement. The beauty and simplicity of mathematical morphology approach come from the fact that a large class of filters can be represented as the combination of two simple operations: erosion and dilation. Let Z denote the set of integers and $f(i, j)$ denote a discrete image signal, where the domain set is given by

$\{i, j\} \in N_1 \times N_2$, $N_1 \times N_2 \subset Z^2$ and the range set by $\{f\} \in N_3$, $N_3 \subset Z$. A structuring element B is a subset in Z^2 with a simple geometrical shape and size. Denote $B^s = \{-b : b \in B\}$ as the symmetric set of B and B_{t_1,t_2} as the translation of B by (t_1, t_2), where $(t_1, t_2) \in Z^2$. The erosion $f \ominus B^s$ and dilation $f \oplus B^s$ can be expressed as [19]

$$(f \ominus B^s)(i, j) = \min_{t_1,t_2 \in B_{i,j}} (f(t_1, t_2)), \tag{1}$$

$$(f \oplus B^s)(i, j) = \max_{t_1,t_2 \in B_{i,j}} (f(t_1, t_2)). \tag{2}$$

On the other hand, opening $f \circ B$ and closing $f \bullet B$ are defined as [19]

$$(f \circ B)(i, j) = ((f \ominus B^s) \oplus B)(i, j), \tag{3}$$

$$(f \bullet B)(i, j) = ((f \oplus B^s) \ominus B)(i, j). \tag{4}$$

A gray value image can be viewed as a two-dimensional surface in a three-dimensional space. Given an image, the opening operation removes the objects, which have size smaller than the structuring element, with positive intensity. Thus, with the specified structuring element, one can extract different image contexts by taking the difference between the original and opening processed image, which is known as "tophat" operation [19].

II.2   Morphological Enhancement Algorithms

Based on the properties of morphological filters, we designed one type of mass pattern-dependent enhancement approaches. The algorithm is implemented by dual morphological tophat operations following by a subtraction which is described as follows.

Step 1: The textures without the pattern information of interest are extracted by a tophat operation

$$r_1(i, j) = \max(0, [f(i, j) - (f \circ B_1)(i, j)]) \tag{5}$$

where $f(i, j)$ is the original image, and $r_1(i, j)$ is the residue image between the original image and the opening of the original image by a specified structuring element B_1. The size of B_1 should be chosen smaller than the size of masses.

Step 2: Let $r_2(i,j)$ be the mass pattern enhanced image by background correction, i.e., by the second tophat operation on $f(i,j)$:

$$r_2(i,j) = \max(0, [f(i,j) - (f \circ B_2)(i,j)]). \tag{6}$$

where B_2 is a specified structuring element which has a larger size than masses.

Step 3: The enhanced image $f_1(i,j)$ can be derived as

$$f_1(i,j) = \max(0, [r_2(i,j) - r_1(i,j)]). \tag{7}$$

This operation is called "dual morphological operation". It can remove the background noise and the structure noise inside the suspected mass patterns. Fig. 2 shows the mass patch and the enhanced results of each step using the dual morphological operation. As we can see from Fig. 2, both background correction (Fig. 2 (c)) and dual morphological operation (Fig. 2 (d)) enhanced the mass pattern, but dual morphological operation removed more structural noise inside the mass region which in turn would improve the mass segmentation results.

III Model-Based Segmentation

III.1 Statistical Modeling

Given a digital image consisting of $N_1 \times N_2$ pixels, assume this image contains K regions. By randomly reordering all pixels in the underlying probability space, one can treat pixel labels as random variables and introduce a prior probability measure π_k. Then the FGGM probability density function (pdf) of gray-level of each pixel is given by [17]:

$$p(x_i) = \sum_{k=1}^{K} \pi_k p_k(x_i), \quad i = 1, \cdots, N_1 N_2, \quad x_i = 0, 1, \cdots, L-1 \tag{8}$$

where x_i is the gray-level of pixel i, and L is the number of gray levels. $p_k(x_i)$'s are conditional region pdf's with the weighting factor π_k, satisfying $\pi_k > 0$, and $\sum_{k=1}^{K} \pi_k = 1$. The generalized Gaussian pdf given region k is defined by

$$p_k(x_i) = \frac{\alpha \beta_k}{2\Gamma(1/\alpha)} \exp\left[-|\beta_k(x_i - \mu_k)|^\alpha\right], \quad \alpha > 0, \quad \beta_k = \frac{1}{\sigma_k}\left[\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right]^{1/2}. \tag{9}$$

where μ_k is the mean, $\Gamma(\cdot)$ is the Gamma function. β_k is a parameter related to the variance σ_k. It can be shown that when $\alpha = 2.0$, one has the Gaussian pdf; when $\alpha = 1.0$, one has

the Laplacian pdf. When $\alpha \gg 1$, the distribution tends to a uniform pdf; when $\alpha < 1$, the pdf becomes sharp. Therefore, the generalized Gaussian model is a suitable model to fit the histogram distribution of those images whose statistical properties are unknown since the kernel shape can be controlled by selecting different α values.

The whole image can be well approximated by an independent and identically distributed random field $\mathbf{X}$. The corresponding joint pdf is

$$P(\mathbf{x}) = \prod_{i=1}^{N_1 N_2} \sum_{k=1}^{K} \pi_k p_k(x_i) \tag{10}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_{N_1 N_2}]$, and $\mathbf{x} \in \mathbf{X}$. $p_k(x_i)$ is given in (9). Based on the joint probability measure of pixel images, the likelihood function under FGGM modeling can be expressed as $\mathcal{L}(\mathbf{r}) = \prod_{i=1}^{N_1 N_2} p_{\mathbf{r}}(x_i)$ where $\mathbf{r} : \{K, \alpha, \pi_k, \mu_k, \sigma_k, k = 1, \cdots, K\}$ denotes the model parameter set.

III.2 Model Identification

With an appropriate system likelihood function, the objective of model identification is to estimate the model parameters by maximizing the likelihood function, or equivalently minimizing the relative entropy between the image histogram $p_{\mathbf{x}}(u)$ and the estimated pdf $p_{\mathbf{r}}(u)$, where u is the gray level. Based on the FGGM model, the expectation-maximization (EM) algorithm is applied to estimate the model parameters. The EM algorithm is an iterative technique for maximum likelihood estimation [20]. Recently, it has been used in many medical imaging applications [15]. Instead of evaluating directly the value of maximum likelihood, we use the global relative entropy (GRE) between the histogram and the estimated FGGM distribution to measure the performance of parameter estimation, given by

$$GRE(p_{\mathbf{x}} \| p_{\mathbf{r}}) = \sum_{u} p_{\mathbf{x}}(u) \log \frac{p_{\mathbf{x}}(u)}{p_{\mathbf{r}}(u)}. \tag{11}$$

Motivated by the same spirit of conventional EM algorithm for finite normal mixtures, we formulated the EM algorithm to estimate the parameter values of the FGGM. The algorithm is summarized as follows.

EM Algorithm:

1. For $\alpha = \alpha_{min}, \cdots, \alpha_{max}$

   - $m = 0$, given initialized $\mathbf{r}^{(0)}$

   - E-step: for $i = 1, ..., N_1N_2, \; k = 1, ..., K$, compute the probabilistic membership

$$z_{ik}^{(m)} = \frac{\pi_k^{(m)} p_k(x_i)}{\sum_{k=1}^{K} \pi_k^{(m)} p_k(x_i)} \qquad (12)$$

   - M-step: for $k = 1, ..., K$, compute the updated parameter estimates

$$\begin{cases} \pi_k^{(m+1)} = \frac{1}{N_1N_2} \sum_{i=1}^{N_1N_2} z_{ik}^{(m)} \\ \mu_k^{(m+1)} = \frac{1}{N_1N_2\pi_k^{(m+1)}} \sum_{i=1}^{N_1N_2} z_{ik}^{(m)} x_i \\ \sigma_k^{2(m+1)} = \frac{1}{N_1N_2\pi_k^{(m+1)}} \sum_{i=1}^{N_1N_2} z_{ik}^{(m)} (x_i - \mu_k^{(m+1)})^2 \end{cases} \qquad (13)$$

   - When $|GRE^{(m)}(p_{\mathbf{x}}||p_{\mathbf{r}}) - GRE^{(m+1)}(p_{\mathbf{x}}||p_{\mathbf{r}})| \leq \epsilon$ is satisfied, go to Step 2

     Otherwise, $m = m + 1$ and go to E-Step.

2. Compute GRE, and go to Step 1

3. Choose the optimal $\hat{\mathbf{r}}$ which corresponds to the minimum GRE.

As we mentioned in Section I, the two important parameters in model selection are K and α. Determination of the region parameter K directly affects the quality of the resulting model parameter estimation and in turn, affects the result of segmentation. In this paper we propose an approach to determine the value of K based on two popular information theoretic criteria introduced by Akaike [23] and by Rissanen [24]. Akaike proposed to select the model that gives the minimum Akaike Information Criterion (AIC), defined by

$$AIC(K) = -2\log(\mathcal{L}(\hat{\mathbf{r}}_{ML})) + 2K' \qquad (14)$$

where $\hat{\mathbf{r}}_{ML}$ is the maximum likelihood estimate of the model parameter set $\mathbf{r}$, and K' is the number of free adjustable parameters in the model [15, 23]. AIC criterion will select the correct number of the image regions K_0 when

$$K_0 = \arg \left\{ \min_{1 \leq K \leq K_{MAX}} AIC(K) \right\}. \qquad (15)$$

Rissanen addressed the problem from a quite different point of view. Rissanen reformulated the problem explicitly as an information coding problem in which the best model fitness is measured such that it assigns high probabilities to the observed data while at the same time

the model itself is not too complex to describe [24]. The model is selected by minimizing the total description length defined by

$$MDL(K) = -\log(\mathcal{L}(\hat{\mathbf{r}}_{ML})) + 0.5K'\log(N_1 N_2). \tag{16}$$

Similarly, the correct number of the distinctive image regions K_0 will be estimated when

$$K_0 = \arg\left\{\min_{1 \leq K \leq K_{MAX}} MDL(K)\right\}. \tag{17}$$

III.3 Bayesian Relaxation Labeling

Once the FGGM model is given, a segmentation problem is the assignment of labels to each pixel in the image. A straightforward way is to label pixels into different regions by maximizing the individual likelihood function $p_k(x)$. This approach is called ML classifier, which is equivalent to a multiple thresholding method. Usually, this method may not achieve a good performance since there is lack of local neighborhood information to be included to make a good decision. CBRL algorithm [25] is one of the approaches, which can incorporate the local neighborhood information into labeling procedure and thus improve the segmentation performance. In this study, we developed the CBRL algorithm to perform/refine pixel labeling based on the localized FGGM model, which is defined as follows:

Let ∂i be the neighborhood of pixel i with an $m \times m$ template centered at pixel i. An indicator function is used to represent the local neighborhood constraints $R_{ij}(l_i, l_j) = I(l_i, l_j)$, where l_i and l_j are labels of pixels i and j, respectively. Note that pairs of labels are now either compatible or incompatible. Similar to reference [25], one can compute the frequency of neighbors of pixel i which has the same label values k as at pixel i

$$\pi_k^{(i)} = p(l_i = k | \mathbf{l}_{\partial i}) = \frac{1}{m^2 - 1} \sum_{j \in \partial i, j \neq i} I(k, l_j) \tag{18}$$

where $\mathbf{l}_{\partial i}$ denotes the labels of the neighbors of pixel i. Since $\pi_k^{(i)}$ is a conditional probability of a region, the localized FGGM pdf of gray-level x_i at pixel i is given by

$$p(x_i | \mathbf{l}_{\partial i}) = \sum_{k=1}^{K} \pi_k^{(i)} p_k(x_i) \tag{19}$$

where $p_k(x_i)$ is given in (9). Assuming gray values of the image are conditional independent, the joint pdf of $\mathbf{x}$, given the context labels $\mathbf{l}$, is

$$P(\mathbf{x}|\mathbf{l}) = \prod_{i=1}^{N_1 N_2} \sum_{k=1}^{K} \pi_k^{(i)} p_k(x_i) \tag{20}$$

where $\mathbf{l} = (l_i : i = 1, \cdots, N_1 N_2)$.

It is known that CBRL algorithm can obtain a consistent labeling solution based on the localized FGGM model (19). Since $\mathbf{l}$ represents the labeled image, it is consistent if $S_i(l_i) \geq S_i(k)$, for all $k = 1, \cdots, K$ and for $i = 1, \cdots, N_1 N_2$ [25], where

$$S_i(k) = \pi_k^{(i)} p_k(x_i). \tag{21}$$

Now we can define

$$A(\mathbf{l}) = \sum_{i=1}^{N_1 N_2} \left(\sum_{k} I(l_i, k) S_i(k) \right) \tag{22}$$

as the average measure of local consistency, and

$$LC_i = \sum_{k} I(l_i, k) S_i(k), \quad i = 1, \cdots, N_1 N_2 \tag{23}$$

represents the local consistency based on $\mathbf{l}$. The goal is to find a consistent labeling $\mathbf{l}$ which can maximize (22). In the real application, each local consistency measure LC_i can be maximized independently. In [25], it has been shown that when $R_{ij}(l_i, l_j) = R_{ji}(l_j, l_i)$, if $A(\mathbf{l})$ attains a local maximum at $\mathbf{l}$, then $\mathbf{l}$ is a consistent labeling.

Based on the localized FGGM model, $l_i^{(0)}$ can be initialized by ML classifier,

$$l_i^{(0)} = \arg\{\max_{k} \ p_k(x_i)\}, \quad k = 1, \cdots, K. \tag{24}$$

Then, the order of pixels is randomly permutated and each label l_i is updated to maximize LC_i, i.e., classify pixel i into kth region if

$$l_i = \arg\{\max_{k} \ \pi_k^{(i)} p_k(x_i)\}, \quad k = 1, \cdots, K \tag{25}$$

where $p_k(x_i)$ is given in (9), $\pi_k^{(i)}$ is given in (18). By considering (24) and (25), we developed a modified CBRL algorithm as follows:

CBRL Algorithm:

1. Given $l^{(0)}$, m=0

2. Update pixel labels

   - Randomly visit each pixel for $i = 1, .., N_1 N_2$

   - Update its label l_i according to

$$l_i^{(m)} = \arg \left\{ \max_k \pi_k^{(i)(m)} p_k(x_i) \right\}$$

3. When $\frac{\sum (\mathbf{l}^{(m+1)} \oplus \mathbf{l}^{(m)})}{N_1 N_2} \leq 1\%$, stop; otherwise, $m = m + 1$, and repeat Step 2.

IV   Experimental Results and Discussion

In this section, we present the results of using the morphological filtering and model-based segmentation approach we have introduced for enhancement and segmentation of suspicious masses in mammographic images. In addition to the qualitative assessment by the radiologists, we introduce several objective measures to assess the performance of the algorithms we have proposed for enhancement and segmentation.

A testing data set of 200 mammograms and two simulated tone images were used to test and evaluate the performance of the algorithms in this study. The mammograms were selected from the Mammographic Image Analysis Society (MIAS) database and the Brook Army Medical Center (BAMC) database created by the Department of Radiology at Georgetown University Medical Center. Of the 200 mammograms, 50 mammograms are normal, and each of the 150 abnormal mammograms contains at least one mass case of varying size, subtlety, and location. The areas of suspicious masses were identified by an expert radiologist based on visual criteria and biopsy proven results. The total data set includes 113 benign and 73 malignant masses. The distribution of the masses in terms of size is shown in Table 1. The BAMC films were digitized with a laser film digitizer (Lumiscan 150) at a pixel size of $100\mu m \times 100\mu m$ and 4096 gray levels (12 bits). Before the method was applied the digital mammograms were smoothed by averaging 4×4 pixels into one pixel. According to radiologists, the size of small masses is $3 - 15mm$ in effective diameter. A $3mm$ object in an original mammogram occupies 30 pixels in a digitized image with a $100\mu m$ resolution. After reducing the image size by four times, the object will occupy the range of about $7 - 8$ pixels. The object with the size of 7

pixels is expected to be detectable by any computer algorithm. Therefore, the shrinking step is applicable for mass cases and can save computation time.

Experimental Evaluation of Morphological Enhancement

In order to justify the suitability of morphological structural elements, the geometric properties of the contexts and textures in mammograms were studied. The basic idea is to keep all mass-like objects within certain size range and remove all others by using the proposed morphological filters with specific structural elements. At the resolution of $400\mu m$, a disk with a diameter of 7 pixels was chosen as the morphological structuring elements B_1 to extract textures in mammograms. Since the smallest masses have 7 pixels in diameter with the resolution of $400\mu m$, this procedure would not destroy mass information. For the purpose of background correction, a disk with a diameter of 75 pixels was used as the morphological structuring element B_2. An object with a diameter of 75 pixels corresponds to $30mm$ in the original mammogram. This indicates that all masses with sizes up to $30mm$ can be enhanced by background correction. Masses larger than $30mm$ are rare cases in the clinical setting. In the last stage of our approach, we applied morphological opening and closing filtering using a disk with a diameter of 5 to eliminate small objects which also contribute to texture noise.

All testing mammograms were processed using the proposed enhancement approach with the suggested structuring element B_1 and B_2. Fig. 5 shows processed mammogram examples using the morphological enhancement. Compared the enhanced results (Fig. 5 (b) and (d)) with the original mammograms (Fig. 5(a) and (c)), the proposed method not only enhanced all suspected mass patterns and reduced the texture noise, but also removed the background noise. In summary, the proposed morphological enhancement approach can enhance mass patterns and remove texture structure noises. For dense mammograms, such as the second example in Fig. 5 (c) and (d), the mass is obscured by dense fibroglandular tissues, our experience shows applying the dual morphological operation to remove the fibroglandular tissue background is useful. In addition to the visual evaluation by the radiologist, we performed the segmentation to assess the effectiveness of the morphological filtering, based on the enhanced mammograms and the original mammograms.

Simulated Evaluation of Segmentation Algorithms

The performance of model selection using two frequently-used methods, i.e., the AIC and MDL [22], were first tested and compared in the simulation study. The computer-generated data was made up of four overlapping normal components. Each component represents one local region. The value for each component were set to a constant value, the noise of normal distribution was then added to this simulation digital phantom. Three noise levels with different variance were set to keep the same signal-to-noise ratio (SNR), where SNR is defined by

$$SNR = 10 \log_{10} \frac{(\Delta \mu)^2}{\sigma^2} \tag{26}$$

where $\Delta \mu$ is the mean difference between regions, and σ^2 is the noise power. The original data for the simulation study are given in Fig. 3 (a). The AIC and MDL curves, as functions of the number of local clusters K, are plotted in Fig. 3 (b). According to the information theoretic criteria, the minima of these curves indicate the correct number of the local regions. From this experimental figure, it is clear that the number of local regions suggested by these criteria are all correct.

For the validation of image segmentation using CBRL, we apply the algorithm first to a simulated image. We use ML classifier to initialize image segmentation, i.e., to initialize the quantified image by selecting the pixel label with largest likelihood at each node. The classification error after initialization is uniformly distributed over the spatial domain as shown in Fig. 4 (a). Our experience suggested this to be a very suitable starting point for contextual relaxation labeling [21]. The CBRL is then performed to fine tune the image segmentation. It should be emphasized that the ground truth is known in this simulated experiment, the percentage of total classification error is used as the criterion for evaluating the performance of segmentation technique. In Fig. 4 (a)-(d), the initial segmentation by the ML classification and the stepwise results of three iterations in the CBRL are presented. In this experiment, algorithm initialization results in an average classification error of 30%. It can be clearly seen that a dramatic improvement is obtained after several iterations of the CBRL by using local constraints determined by the context information. In addition, the convergence is fast as one can see, after the first iteration most of the misclassification are removed. We have also implemented two other independent and popular algorithms, namely, the iterated conditional

mode (ICM) and the modified iterated conditional mode (MICM) algorithms, so as to assess the comparative performance of the segmentation results among different approaches [21, 22]. The only assumption being made by these three methods is the Markovian property of the context images which can be well justified by the underlying cell oncology and pathology. We have applied these three algorithms to the same testing image and the corresponding classification errors are presented in Table 2. The final percentage of classification errors for Fig. 4 (d) is 0.7935%. From this experimental comparison, it can be concluded that three algorithms achieved comparable segmentation accuracy and the result produced by the MICM algorithm is most superior, though in terms of computational complexity the CBRL algorithm is the least. It should be noticed that since in MICM algorithm an inhomogeneous configuration of the Markov random field is used, its superior performance is reasonable.

On Model-Based Segmentation - Real Case Study

In the real case study, we used two information criteria (AIC and MDL) to determine K. Table 3 and Table 4 shows the AIC and MDL values with different K and α of the FGGM model based on one original mammogram. As it can be seen from Table 3 and Table 4, although with different α, all AIC and MDL values achieve the minimum when $K = 8$. It indicates that AIC and MDL are relatively insensitive to the change of α. With this observation, we can decouple the relation between K and α and choose the appropriate value of one while fixing the value of another. Fig. 6 (a) and Fig. 6 (b) are two examples of AIC and MDL curves with different K and fixed $\alpha = 3.0$. Fig. 6 (a) is based on the original mammogram and Fig. 6 (b) is based on the enhanced mammogram. As we can see in Fig. 6 (a), both criteria achieved the minimum when $K = 8$. It should be noticed that though no ground truth is available in this case, our extensive numerical experiments have shown a very consistent performance of the model selection procedure and all the conclusions were strongly supported by the previous independent work reported by [14]. Fig. 6 (b) indicates that $K = 4$ is the appropriate choice for the mammogram enhanced by dual morphological operation. This is believed to be reasonable since the number of regions decrease after background correction.

We fixed $K = 8$, and changed the value of α for estimating the FGGM model parameters using the proposed EM algorithm with the original mammogram The GRE value between the

histogram and the estimated FGGM distribution was used as a measure of the estimation bias. We found that GRE achieved a minimum distance when the FGGM parameter $\alpha = 3.0$ as shown in Fig. 7. The similar result was shown when we applied the EM algorithm to the enhanced mammogram with $K = 4$. This indicated that the FGGM model might be better than the finite normal mixture (FNM) model ($\alpha = 2.0$) in modeling mammographic images when the true statistical properties of mammograms are generally unknown, though the FNM has been most often chosen in many previous work [15].

After the determination of all model parameters, every pixel of the image was labeled to different region (from 1 to K) based on the CBRL algorithm. We then selected the brightest region, which corresponding to label K, plus a criterion of closed isolated area, as the candidate region of suspicious masses. According to the visual inspections by the radiologists, when we use $K - 1$ instead of K, the results are over-segmented. For the case of using $K + 1$, the results are under-segmented. In order to quantify the performance differences between the different segmentation methods, several groups have suggested that the segmentation results may be compared against radiologists' outlines of the lesions [3]. Though the proposed comparison measures are quantitative, the performance measures are still qualitative, since the reference base (e.g., gold standard by the radiologists) is qualitative, subjective, and imperfect. Therefore, in this model-supported approach, in addition to the visual inspections by the radiologists, we have also introduced an objective measure, the global relative entropy between the histogram of the pixel images $p_{\mathbf{x}}(u)$ and the FGGM of the *segmented* image $p_{\mathbf{x},l}(u)$ to assess the performance of the segmentation, defined by

$$GRE(p_{\mathbf{x}}(u)\|p_{\mathbf{x},l}(u)) = \sum_u p_{\mathbf{x}}(u) \log \frac{p_{\mathbf{x}}(u)}{p_{\mathbf{x},l}(u))} \tag{27}$$

where l is the context image estimated by the segmentation algorithm. Considering that the ergodic theorem is the most fundamental principle in the detection and estimation theory, it is believed that when a good segmentation is achieved, the distance between the $p_{\mathbf{x}}(u)$ and $p_{\mathbf{x},l}(u)$ should be minimized and this measure links the image text and its sample averages. Our experience has suggested that this post-segmentation measure may be a suitable objective criterion for evaluating the quality of image segmentation in a fully unsupervised situation [22, 26, 27, 28]. Table 5 shows our evaluation data from three different segmentation methods

when applied to the real images.

Performance of Combined Morphological Filtering and Model-Based Segmentation using a Larger Database

The proposed segmentation method was used to extract suspicious mass regions from the 200 testing mammograms. Without enhancement, a total of 1142 potential mass regions were isolated including 114 of the 186 true masses. With enhancement, a total of 3143 potential mass regions were extracted including 181 of the 186 true masses. The results demonstrated that more true masses were picked up after enhancement although more false cases were also included. The undetected areas mainly occurred at the lower intensity side of the shaded objects or obscured by fibroglandular tissues that, however, were extracted on morphological enhanced mammograms. In addition, when the margins of masses are ill defined, only parts of suspicious masses were extracted from the original mammograms. For the purpose of "lesion site selection", we believe that the sensitivity should be the sole criterion for the performance evaluation of the method. We have 181/186 vs. 114/186. Our method is unsupervised and automatic and does not involve any detection effort at this moment. To our best knowledge, there is no objective criterion available for the evaluation of image enhancement performance before a detection effort is involved. We only claimed that the enhancement step is important and effective with respect to the purpose of "lesion site selection".

Fig. 8 demonstrates some segmentation results based on the original and enhanced mammograms. We compared the segmentation results based on the enhanced mammogram ($K = 4$, and $\alpha = 3.0$) with those based on the original mammogram ($K = 8$, and $\alpha = 3.0$) as shown in Fig. 8. Comparing the results in Fig. 8 (b) with those in Fig. 8 (a), we can see that after enhancement, a more accurate region was detected for the suspected mass which has ill-defined margin. Getting an accurate suspected region is a crucial issue since geometric features are extracted based on suspected regions and these features are very important for further true mass detection. In addition, we observed that one suspected mass was missed in Fig. 8 (a) but was detected in Fig. 8 (b). As we have mentioned in Section I, none of the suspected masses should be missed in the segmentation step. Fig. 8 (c) and (d) demonstrate the segmentation of a suspected mass that lies in dense breast tissue. As shown in Fig. 8 (c), the whole fibroglan-

dular tissue area was segmented when based on the original mammogram. After enhancement, the suspected region was segmented exactly as shown in Fig. 8 (d).

We have also included the segmentation results on the normal mammograms. Fig. 9 demonstrate the segmentation results based on the original and enhanced mixed fatty and glandular mammograms. Fig. 10 demonstrate the segmentation results based on the original and enhanced dense mammograms. We would like to emphasize that the objective of this paper is to provide a segmentation technique which can enhance and extract potential mass site from the background so that the characterization of the related mass pattern can be accurately extracted in terms of focused feature selection and analysis. The method of course will produce many mass-like areas, but it will be a plausible outcome since the accurate description of non-mass cases characterized by mass-like sites will benefit the follow-on detection step where the performance of the classifier depends on an accurate separation of mass and non-mass in the featured spaces. The details will be described in the Part II of this paper.

For the purpose of evaluating the performance of the segmentation method, we used both simulated studies and expert visual inspection to validate the methods and results. The radiologist has concluded that the lesion characteristics after the proposed enhancement have been better displayed and all possible lesion areas have been successfully identified. In addition to the visual inspection, we have measured the overlap between the computer-segmented and the radiologist segmented mass regions to evaluate our method. Fig. 11 shows the comparison results of segmentation based on the enhanced mammograms. Fig. 11 includes 60 benign and malignant mass patches which were cut from the whole mammograms after the segmentation. The white outline was drawn by the radiologist while the black outline was produced by the computer and was superimposed upon the original image. As we can see from Fig. 11, for most of cases, the ratio of mutual overlap area of the radiologist segmented mass region and the computer-segmented mass region to the radiologist segmented mass area is large than 50%. In addition, even the poorest result picked the true lesion in the correct location and depicted the characteristics of the mass reasonably. It is important to understand that "lesion area segmentation" is not our objective, so there is no "best" or "worst" segmentation results. Our objective is "lesion site selection" with a possible highest sensitivity through a global

unsupervised enhancement and segmentation scheme.

V Conclusion

In this paper, we propose a combined method of using morphological operations, a finite generalized Gaussian mixture modeling, and a contextual Bayesian relaxation labeling to enhance and segment various breast tissue textures and suspicious mass lesions from mammographic images. This phase is a crucial step in mass detection for an improved CAD. We emphasized the importance of model selection which includes the selection of the number of image regions K and the selection of FGGM kernel shape controlled by α. The experimental results indicate that the suspected masse sites selection can be affected by different K and α. We proposed the EM algorithm together with the information theoretic criteria to determine the optimal K and α. With optimal K and α, the segmentation results can be significantly improved. We also showed that with the proposed pattern-dependent enhancement algorithm using morphological operations, the subtle masses can be segmented more accurately than those when the original image is used for extraction without enhancement. To summarize, the morphological filtering enhancement combined with the stochastic model-based segmentation is an effective way to extract mammographic suspicious patterns of interest, and thereby may facilitate the overall performance of mammographic computer-aided diagnosis of breast cancer.

Acknowledgement

References

[1] H. Kobatake, M. Murakami, H. Takeo, and and S. Nawano, "Computerized Detection of Malignant Tumors on Digital Mammograms," *IEEE Trans. on Med. Imaging,* Vol. 18, No. 5, pp. 369-378, 1999.

[2] R. Zwiggelaar, T. C. Parr, J. E. Schumm, I. W. Hutt, C. J. Taylor, S. M. Astley, and C. R. M. Boggis, "Model-based Detection of Spiculated Lesions in Mammograms," *Medical Image Analysis,* Vol. 3, No. 1, pp. 39-62, 1999.

[3] M. A. Kupinski and M. L. Giger, "Automated Seeded Lesion Segmentation on Digital Mammograms," *IEEE Trans. Medical Imaging,* vol. 17, No. 4, pp. 510-517, 1998.

[4] N. Karssemeijer and G. M. te Brake, "Detection of Stellate Distortions in Mammogram," *IEEE Trans. on Med. Imaging,* Vol. 15, pp. 611-619, 1996.

[5] W. K. Zouras, M. L. Giger, P. Lu, D. E. Wolverton, C. J. Vyborny, and K. Doi, "Investigation of a Temporal Subtraction Scheme for Computerized Detection of Breast Masses in Mammograms," *Excerpta Medica,* Vol. 1119, pp. 411-415, 1996.

[6] N. Petrick, H. P. Chan, B. Sahiner, and D. Wei, "An Adaptive Density-Weighted Contrast Enhancement Filter for Mammographic Breast Mass Detection," *IEEE Trans. on Med. Imaging,* Vol. 15, No. 1, pp. 59-67, 1996.

[7] M. Sameti and R. K. Ward, "A Fussy Segmentation Algorithm for Mammogram Patttion," *Digital Mammgraphy,* K. Doi, Ed., International Congress Series, Amsterdam, the Netherlands: Elsevier, pp. 471-474, 1996.

[8] W. P. Kegelmeyer Jr., J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided Mammographic Screening for Spiculated Lesions," *Radiology,* Vol. 191, pp. 331-337, 1994.

[9] F. F. Yin, M. L. Giger, C. J. Vyborny, K. Doi, and R. A. Schmidt, "Comparison of Bilateral-Subtraction and Single-Image Processing Techniques in the Computerized Detection of Mammographic Masses," *Investigative Radiology,* Vol. 28, No. 6, pp. 473-481, 1993.

[10] B. Zheng, Y. H. Chang, and D. Gur, "Computerized Detection of Masses in Digitized Mammograms Using Single-Image Segmentation and a Multilayer Topographic Feature Analysis," *Acad. Radiology,* Vol. 2, pp. 959-966, 1995.

[11] H. D. Li, M. Kallergi, L. P. Clarke, V. K. Jain, and R. A. Clark, "Markov Random Field for Tumor Detection in Digital Mammography," *IEEE Trans. on Med. Imaging,* Vol. 14, No. 43 pp. 565-576, 1995.

[12] M. L. Giger, C. J. Vyborny, and R. A. Schmidt, "Computerized Characterization of Mammographic Masses: Analysis of Spiculation," *Cancer Letters,* Vol. 77, pp. 201-211, 1994.

[13] T. K. Lau and W. F. Bischof, "Automated Detection of Breast Tumors Using the Asymmetry Approach," *Computers and Biomed. Res.,* Vol. 24, No. 9, pp. 1501-1513, 1995.

[14] M. J. Bianchi, A. Rios, and M. Kabuka, "An Algorithm for Detection of Masses, Skin Contours, and Enhancement of Microcalcifications in Mammograms," *Proc. , Symposium for Computer Assisted Radiology,* pp. 57-64, Winston-Salem, June 1994.

[15] T. Lei and W. Sewchand, "Statistical Approach to X-Ray CT Imaging and Its Application in Image Analysis–Part II: A New Stochastic Model-Based Image Segmentation Technique for X-Ray CT Image," *IEEE Trans. on Med. Imaging,* Vol. 11, No. 1, pp. 62-69, 1992.

[16] Y. Wang, T. Adali, and S-C. B. Lo, "Automatic Threshold Selection Using Histogram Quantization," *SPIE J. Biomedical Optics,* Vol. 2, No. 2, pp.211-217, April 1997.

[17] J. Zhang and J. W. Modestino, "A Model-Fitting Approach to Cluster Validation with Application to Stochastic Model-Based Image Segmentation," *IEEE Trans. on PAMI,* Vol. 12, No. 10, pp. 1009-1017, 1990.

[18] H. Li, K. J. R. Liu, Y. Wang, and S. C. Lo, "Morphological Filtering and Stochastic Modeling-Based Segmentation of Masses on Mammographic Images," *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference,* pp. 1792-1796, 1996.

[19] J. Serra, *Image Analysis and Mathematical Morphology.* London, U. K.: Academic, 1982.

[20] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. Roy. Statist. Soc. Ser. B*, Vol. 39, pp. 1-38, 1977.

[21] Y. Wang, T. Adali, C. M. Lau, and S. Y Kung, "Quantitative Analysis of MR Brain Image Sequences by Adaptive Self-organizing Finite Mixtures," *J. of VLSI Signal Processing*, Vol. 18, No. 3, pp. 219-240, 1998.

[22] Y. Wang, T. Adali, S. Y. Kung, and Z. Szabo, "Quantification and Segmentation of Brain Tissues from MR Images: A Probabilistic Neural Network Approach," *IEEE Trans. Image Processing*, Vol. 7, No. 8, pp. 1165-1181, 1998.

[23] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans. on Automatic Control*, Vol. 19, No. 6, pp. 716-723, 1974.

[24] J. Rissanen, "Modeling by Shortest Data Description," *Automat.*, Vol. 14, pp. 465-471, 1978.

[25] R. A. Hummel and S. W. Zucker, "On the Foundations of Relaxation Labeling Processes," *IEEE Trans. on PAMI*, Vol. 5, No. 3, pp. 267-286, 1983.

[26] A. Hoover, G. J. Baptoste, X. Jiang, P. J. Flynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, and R. B. Fisher, "An Experimental Comparison of Range Image Segmentation Algorithms," *IEEE Trans. on PAMI*, Vol. 18, No. 7, pp. 673-688, 1996.

[27] Y. J. Zhang, "A Survey on Evaluation Methods for Image Segmentation," *Pattern Recognition*, Vol. 29, No. 8, pp. 1335-1346, 1996.

[28] A. M. Bensaid, L. O. Hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, and R. F. Murtagh, "Validity-Guided Clustering with Applications to Image Segmentation," *IEEE Trans. Fuzzy Systems*, Vol. 4, No. 2, pp. 112-123, 1996.

	$0-5mm$	$6-10mm$	$11-15mm$	$16-20mm$	$21-25mm$	$26-30mm$
#	3	55	78	29	17	4

Table 1: The distribution of the effective size of the 186 masses used in this study. The effective size is defined as the square root of the product of the maximum and minimum diameters of the mass.

Item	CBRL Result	ICM Result	MICM Result
Classification Error	0.7935%	0.7508%	0.3113%

Table 2: Comparison of CBRL, ICM, and MICM Algorithm: Simulated Data.

K	$\alpha = 1.0$	$\alpha = 2.0$	$\alpha = 3.0$	$\alpha = 4.0$
2	651250	650570	650600	650630
3	646220	644770	645280	646200
4	645760	644720	645260	646060
5	645760	644700	645120	646040
6	645740	644670	645110	645990
7	645640	644600	645090	645900
8	645550(min)	644570(min)	645030(min)	645850(min)
9	645580	644590	645080	645880
10	645620	644600	645100	645910

Table 3: Computed AIC's for the FGGM Model with Different α.

K	$\alpha = 1.0$	$\alpha = 2.0$	$\alpha = 3.0$	$\alpha = 4.0$
2	651270	650590	650630	650660
3	646260	644810	645360	646350
4	645860	644770	645280	646150
5	645850	644770	645280	646100
6	645790	644750	645150	646090
7	645720	644700	645120	645930
8	645680(min)	644690(min)	645100(min)	645900(min)
9	645710	644710	645140	645930
10	645790	644750	645180	645960

Table 4: Computed MDL's for the FGGM Model with Different α.

Method	Soft Classification	Bayesian Classification	CBRL
GRE Value	0.0067	0.4406	0.1578

Table 5: Comparison of Segmentation Error Resulting From Noncontextual and Contextual Methods.
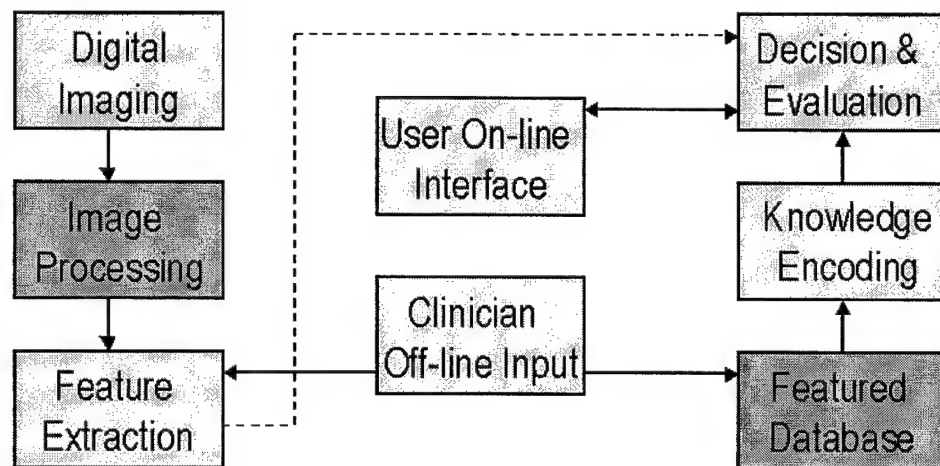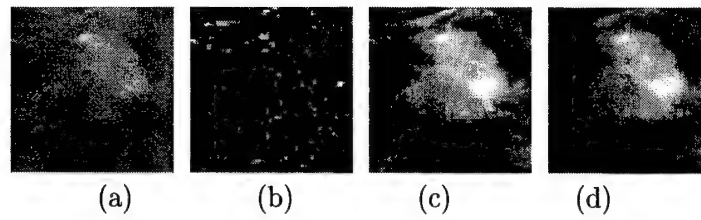
Figure 1: Major components in computer-aided diagnosis.

Figure 2: The original and enhancement result of the mass patch using dual-morphological operation. (a) the original image block $f(i,j)$; (b) the textures $r_1(i,j)$; (c) the background correction result $r_2(i,j)$; (d) the enhanced result $f_1(i,j)$.

(a)   (b)
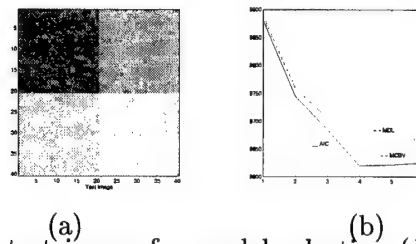
Figure 3: Original simulated test image for model selection ($K_0 = 4$, SNR=10 dB) and the AIC/MDL curves in model selection ($\sigma = 30$).

(a) ML initialization   (b) First iteration in CBRL

(c) Second iteration in CBRL   (d) Third iteration in CBRL
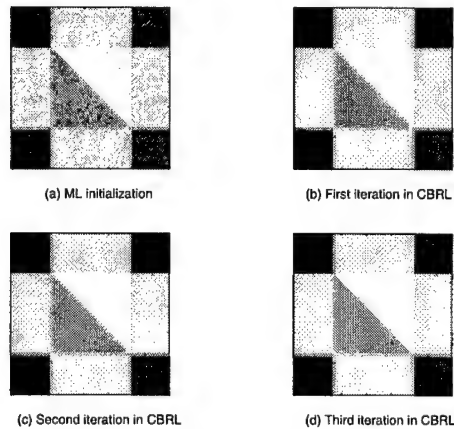
Figure 4: Image segmentation by CBRL on simulated image (with initialization by ML classification).
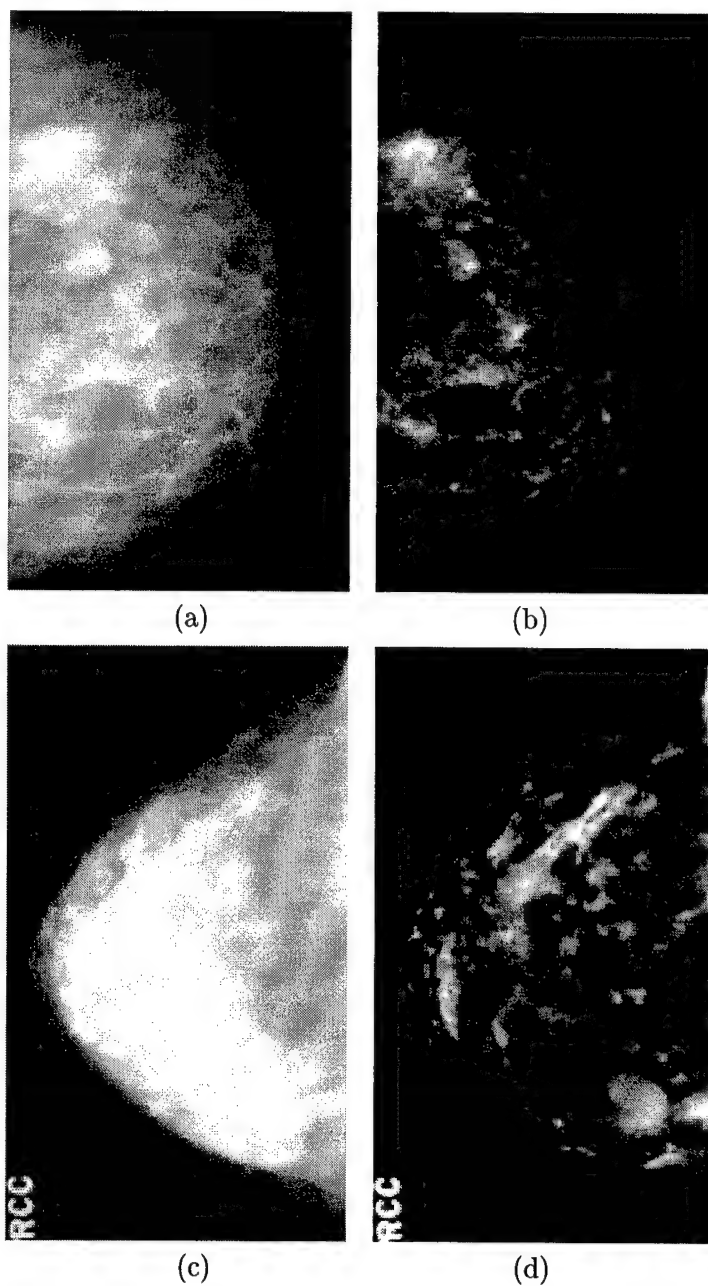
Figure 5: The examples of mass enhancement. (a) original mammogram, (b) enhanced mammogram. (c) and (d) are another original mammogram and its enhanced result.
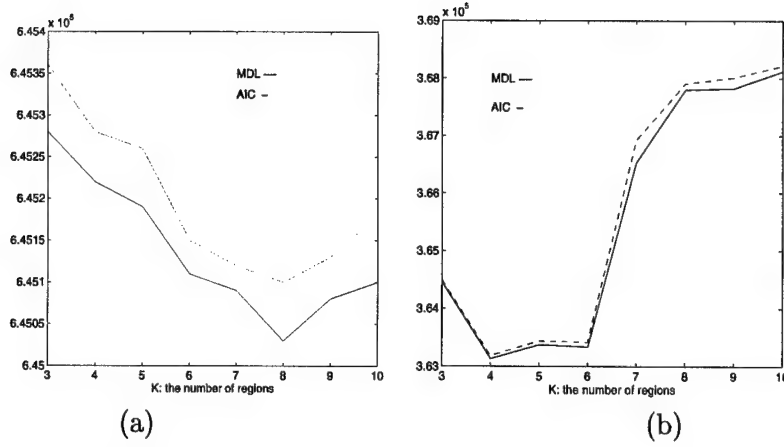
Figure 6: The AIC and MDL curves with different number of region K. (a) the results based on the original mammogram, the optimal $K = 8$; (b) the results based on the enhanced mammogram, the optimal $K = 4$.
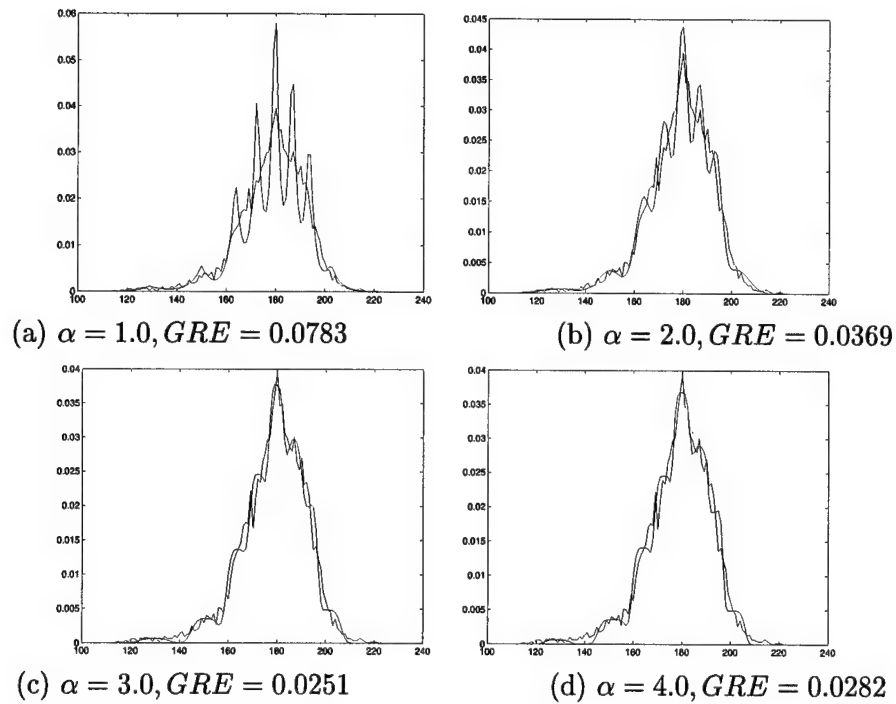

(a) $\alpha = 1.0, GRE = 0.0783$

(b) $\alpha = 2.0, GRE = 0.0369$

(c) $\alpha = 3.0, GRE = 0.0251$

(d) $\alpha = 4.0, GRE = 0.0282$

Figure 7: The comparison of learning curves and histogram of the original mammogram with different α, $K = 8$. The optimal $\alpha = 3.0$.
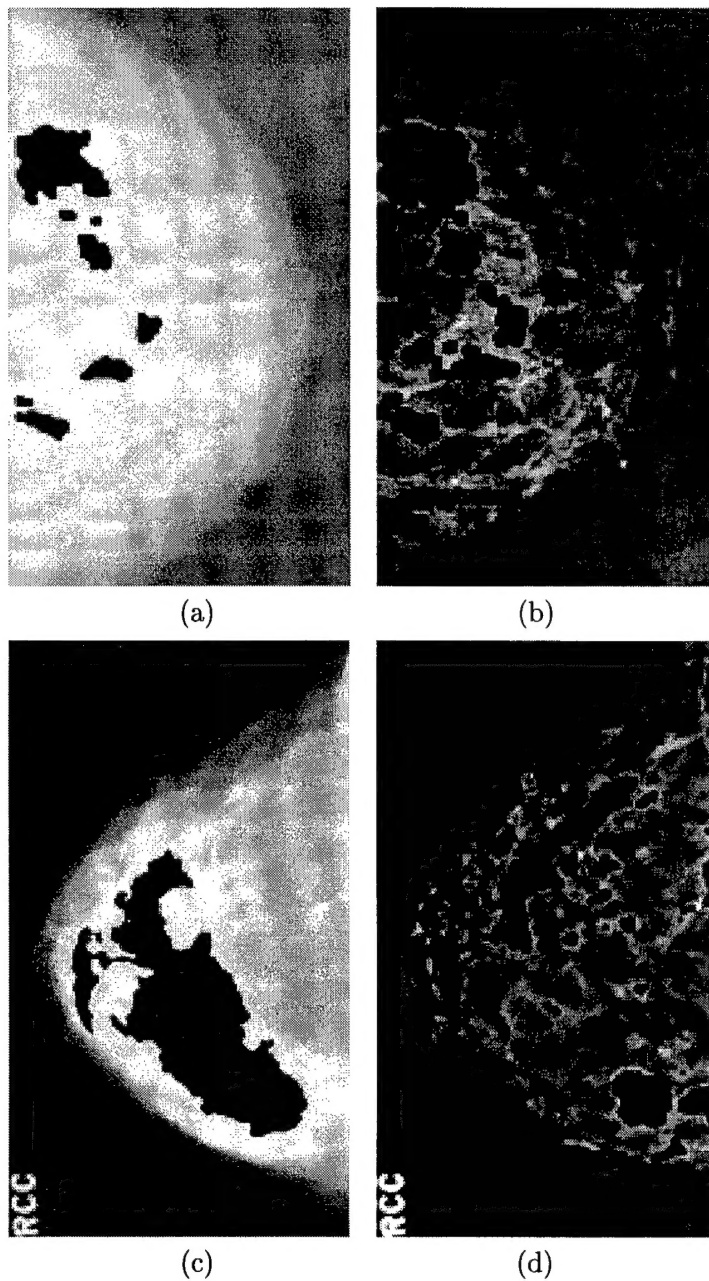
Figure 8: (a) The suspected mass segmentation results based on the original mammogram, (b) the results based on the enhanced mammogram, $K = 4$, $\alpha = 3.0$. (c) and (d) are the results based on another original mammogram and its enhanced image.
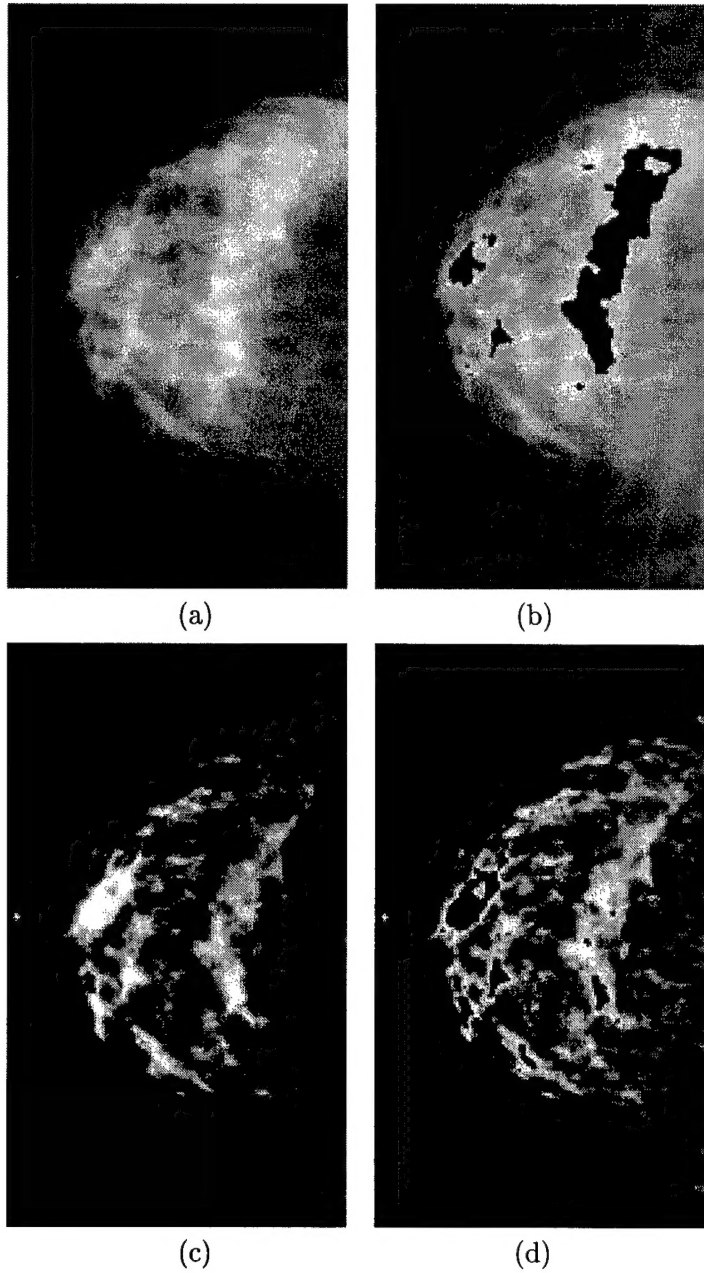
Figure 9: The examples of normal mixed fatty and glandular mammogram. (a) original mammogram, (b) the segmentation result based on the original mammogram, (c) enhanced mammogram, (d) the results based on the enhanced mammogram, $K = 4$, $\alpha = 3.0$.
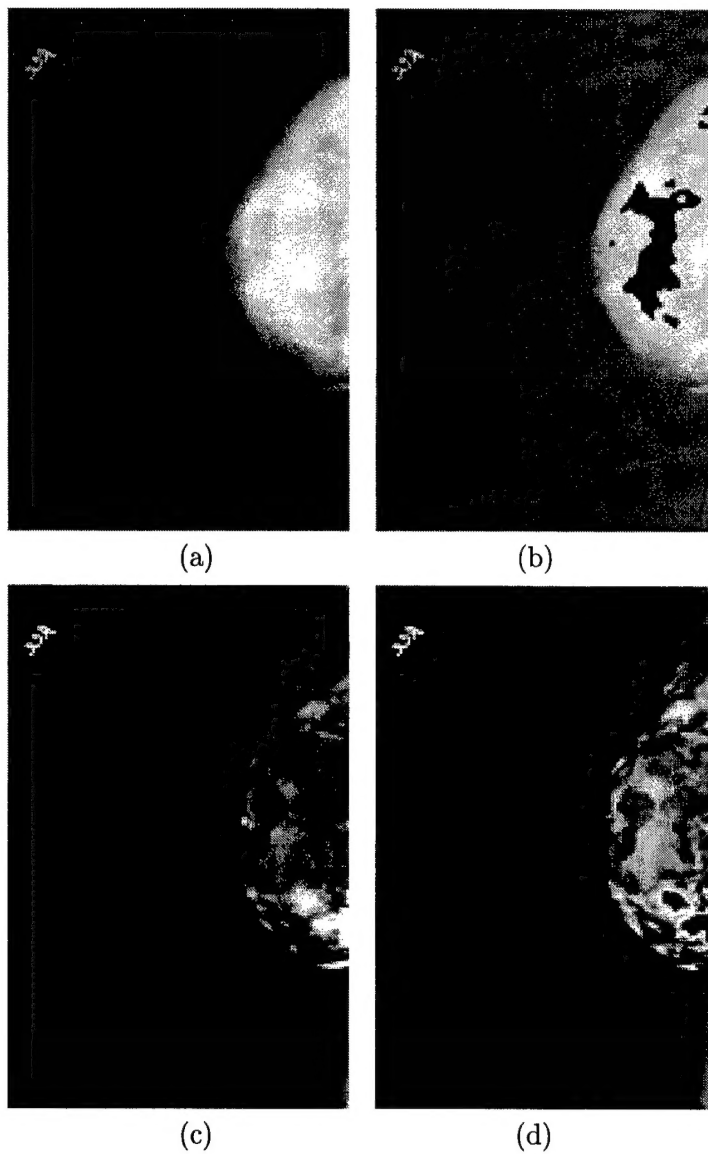
Figure 10: The examples of normal dense mammogram. (a) original mammogram, (b) the segmentation result based on the original mammogram, (c) enhanced mammogram, (d) the results based on the enhanced mammogram, $K = 4$, $\alpha = 3.0$.
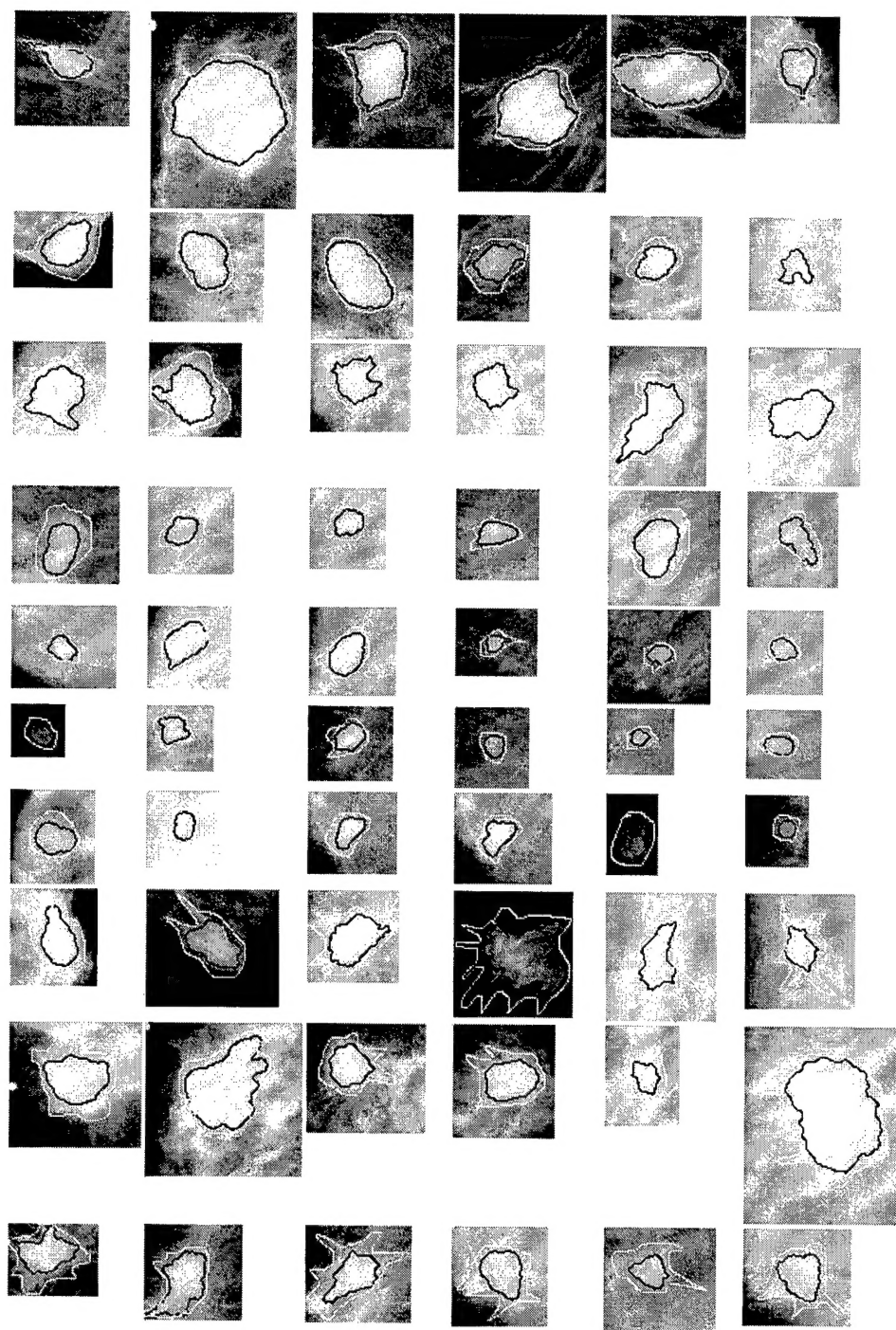
Figure 11: The comparison results of segmentation based on the enhanced mammograms. Black outlines denote the computer-segmented results. White outlines denote the radiologist-segmented results.